

Pulsing response of the cardiac transcriptome

Citation for published version (APA):

Martherus, S. R. M. (2011). *Pulsing response of the cardiac transcriptome*. [Doctoral Thesis, Maastricht University]. Datawyse / Universitaire Pers Maastricht. <https://doi.org/10.26481/dis.20110907sm>

Document status and date:

Published: 01/01/2011

DOI:

[10.26481/dis.20110907sm](https://doi.org/10.26481/dis.20110907sm)

Document Version:

Publisher's PDF, also known as Version of record

Please check the document version of this publication:

- A submitted manuscript is the version of the article upon submission and before peer-review. There can be important differences between the submitted version and the official published version of record. People interested in the research are advised to contact the author for the final version of the publication, or visit the DOI to the publisher's website.
- The final author version and the galley proof are versions of the publication after peer review.
- The final published version features the final layout of the paper including the volume, issue and page numbers.

[Link to publication](#)

General rights

Copyright and moral rights for the publications made accessible in the public portal are retained by the authors and/or other copyright owners and it is a condition of accessing publications that users recognise and abide by the legal requirements associated with these rights.

- Users may download and print one copy of any publication from the public portal for the purpose of private study or research.
- You may not further distribute the material or use it for any profit-making activity or commercial gain
- You may freely distribute the URL identifying the publication in the public portal.

If the publication is distributed under the terms of Article 25fa of the Dutch Copyright Act, indicated by the "Taverne" license above, please follow below link for the End User Agreement:

www.umlib.nl/taverne-license

Take down policy

If you believe that this document breaches copyright please contact us at:

repository@maastrichtuniversity.nl

providing details and we will investigate your claim.

Pulsing response of the cardiac transcriptome

Pulsing response of the cardiac transcriptome
Dissertation, Maastricht University, Maastricht, The Netherlands

ISBN 978 94 6159 077 0

Cover design Ruben Martherus

Lay-out: Ruben Martherus

Printed by: Datawyse – Universitaire Pers Maastricht

© Copyright Ruben Martherus, Maastricht 2011

All rights reserved. No part of this thesis may be reproduced or transmitted in any form or by any means electronic or mechanical, Including photocopying, recording or any information storage or retrieval system without the permission in writing from the author, or when appropriate, from the publishers of the publications.

Pulsing response of the cardiac transcriptome

Proefschrift

ter verkrijging van de graad van doctor aan de Universiteit Maastricht
op gezag van de Rector Magnificus, Prof. mr. G.P.M.F. Mols
volgens het besluit van het College van Decanen
in het openbaar te verdedigen op woensdag 7 september 2011 om
10:00 uur

door

Sarkies Ruben Mattheus Martherus

Geboren op 30 oktober 1977 te Eemnes



Promotores

Prof. dr. H.J.M. Smeets
Prof. dr. J.P.M. Geraedts

Copromotor

Dr. T.A.Y. Ayoubi (Katholieke Universiteit Leuven)

Beoordelingscommissie

Prof. dr. L. de Windt (voorzitter)
Dr. M. van Bilsen
Prof. dr. D.P.V. de Kleijn (Universiteit Utrecht)
Prof. dr. P.A.J. Schrauwen
Prof. dr. W. Van de Ven (Katholieke Universiteit Leuven)

Financial support by the Dutch Heart Foundation for the publication of this thesis is gratefully acknowledged

读万卷书，行万里路

Dú wàn juàn shū, xíng wàn lǐ lù

(Read ten thousand books and walk ten thousand miles)
Chinese Proverb

TABLE OF CONTENTS

List of abbreviations	8
Chapter 1 - General Introduction	11
Chapter 2 - Regulatory network of the cardiomyocyte transcriptome	21
Chapter 3 - Creation and validation of a set of cardiovascular promoter reporter constructs	55
Chapter 4 - Electrical stimulation of primary neonatal rat ventricular cardiomyocytes using pacemakers	73
Chapter 5 - Electrical signals affect the cardiomyocyte transcriptome independently of contraction	81
Chapter 6 - Identification of promoter elements prerequisite for cardiac specificity and the cardiac response to electric stimuli	101
Chapter 7 - Wnt/Frizzled signaling and electrical stimuli affect the cardiomyocyte transcriptome via similar mechanisms	123
Chapter 8 - Functional annotation of heart enriched mitochondrial genes GBAS and CHCHD10 through guilt by association	139
Chapter 9 - General Discussion	155
Summary	167
Samenvatting	171
Dankwoord	177
Curriculum Vitae	179
List of publications	180

List of abbreviations

A7R5	rat aortic smooth muscle cell line
ACTC1	cardiac alpha actin
ANF	atrial natriuretic factor
ATP2A2	ATPase, Ca ²⁺ transporting 2 (cardiac form)
bHLH	basic helix-loop-helix
BNP	brain type natriuretic peptide
C2C12	mouse skeletal myoblast cell line
CHCHD10	coiled-coil-helix-coiled-coil-helix domain containing 10
CNS	conserved non-coding sequence
COX	cytochrome-c-oxidase
CSRP3	cysteine and glycine-rich protein 3
DBTSS	Database of transcription start sites
DNA	deoxyribonucleic acid
DPE	downstream promoter element
FABP3	fatty acid binding protein 3,
FZD	frizzled homolog (Drosophila)
GATA4	GATA binding protein 4
GATA6	GATA binding protein 6
GBAS	glioblastoma amplified sequence
HDAC	histone deacetylase
HOX	homeodomain / homeobox
HSPB7	heat shock 27 kDa protein family member 7
kb	kilo bases
LOXL1	lysyl oxidase-like 1
MEF2	myocyte enhancing factor 2
MYBPC1	myosin binding protein C, slow type
MYH6	myosin heavy chain 6 / alpha myosin heavy chain
MYL2	myosin light chain 2
MYL3	myosin light chain 3
MYL7	myosin light chain 7
MYOCD	myocardin
MyoD	myogenic differentiation factor
MYOG	myogenin
NKE	NKX2-5 response elements
NKX2-5	NK2 transcription factor related, locus 5 (Drosophila)
NPPA	natriuretic peptide precursor, atria type
NPPB	natriuretic peptide precursor, brain type

NRVCM	neonatal rat ventricular cardiomyocyte
PWM	positional weight matrix
RNApolII	RNA polymerase II
RT-QPCR	reverse transcription-quantitative polymerase chain reaction
SP1	SP1 transcription factor / specificity protein 1
SRE	serum response element
SRF	serum Response Factor
SURF1	surfeit 1
TBX5	T-box 5
TCAP	telethonin/ titin capping protein
TCF7L2	transcription factor 7-like 2 (T-cell specific, HMG-box)
TF	transcription factor
TFBS	transcription factor binding site
TIC	transcription initiation complex
TSS	transcription start site
WNT	wingless-type MMTV integration site family

CHAPTER 1

GENERAL INTRODUCTION

Introduction

The human body requires a continuous flow of blood for it to function. Blood supplies the tissues with nutrients and oxygen and carries off waste products for disposal. The force required for transporting blood throughout the body is provided by the continuous contractions of the heart. It is overt that disturbances in the function of this vital organ can be disastrous. Decreased heart function can lead to impairments in the function of other organ systems and can eventually become life threatening.

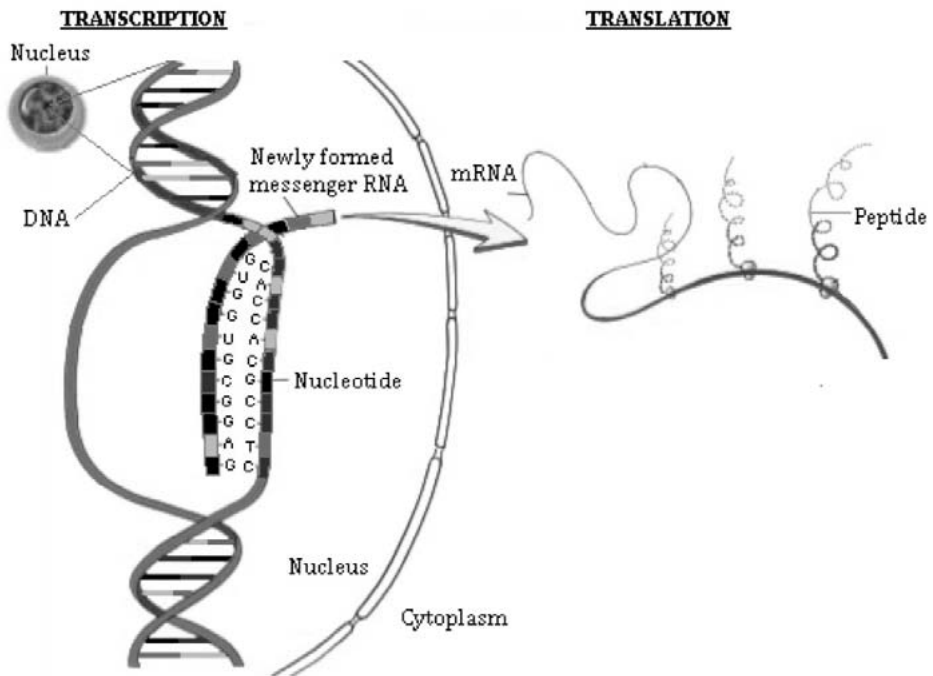


Figure 1. Transcription.

In the nucleus, an mRNA sequence is synthesized based on the genetic code on one of the DNA strands. After transcription, the mRNA is transported to the cytoplasm where the mRNA-sequence is translated into a peptide. This peptide will undergo further processing into its final protein form.

Like any other organ, a healthy functioning heart relies on a tight regulation of the proteome; the integral set of proteins present at any given time and place. The spatial and temporal composition of the proteome is greatly dependent on the composition of the transcriptome, the set of messenger RNA (mRNA) molecules that serve as the template for the synthesis of proteins (translation) (figure 1). A key mechanism in governing the transcriptome is controlling the activation of transcription, the process in which the enzyme RNA-polymerase synthesizes the mRNA intermediate based on its DNA template.

Transcription, interplay between DNA and proteins

Initiation of transcription is the result of a highly complex and tightly regulated interplay between specific sequences located within the DNA and regulatory proteins called transcription factors (TFs). TFs can act via direct binding to the DNA and/or as co-factor by interacting with other TFs. The DNA-binding factors bind to specific regulatory sequences in the DNA, the transcription factor binding sites (TFBs). Most of the regulatory sequences are believed to be located within regions 5' of the genes' coding sequences. For a TF to exert its activity, it usually has to act in concert with others. As a result, co-operating TFBs typically occur together in modules. The DNA-region containing the sequences involved in the transcriptional regulation of a gene is called the promoter (figure 2A). By varying the modular arrangements of TFBs, promoters are formed which are unique for each gene.

Protein-protein and DNA-protein interactions act together in eventually assembling transcription initiating complexes at the specific transcription start sites of genes. To enable the formation of these complexes, some proteins may contribute by opening up the chromatin structures and/or assist in bending the DNA so that co-operating regulatory sequences are brought into close proximity of one another (see figure 2B).[1, 2] In cell-free systems, RNA-polymerase only requires a small set of often ubiquitously expressed transcription factors, such as the TATA-box binding protein, to initiate transcription at basal levels.[3] The promoter region targeted by this set of factors is the core promoter. However, for higher levels of transcription and a spatial and temporal specific transcription, additional factors are required.[4] Which TFs interact with the DNA depends on the availability of TFs within the nucleus of a given cell at a given time and on which TFBs are available on the DNA.

The composition of the complete set of factors involved is highly dynamic. In response to any kind of signal, proteins with a TF-function can become differentially regulated and transmit the signal to the nucleus, resulting in alterations in the transcriptome.

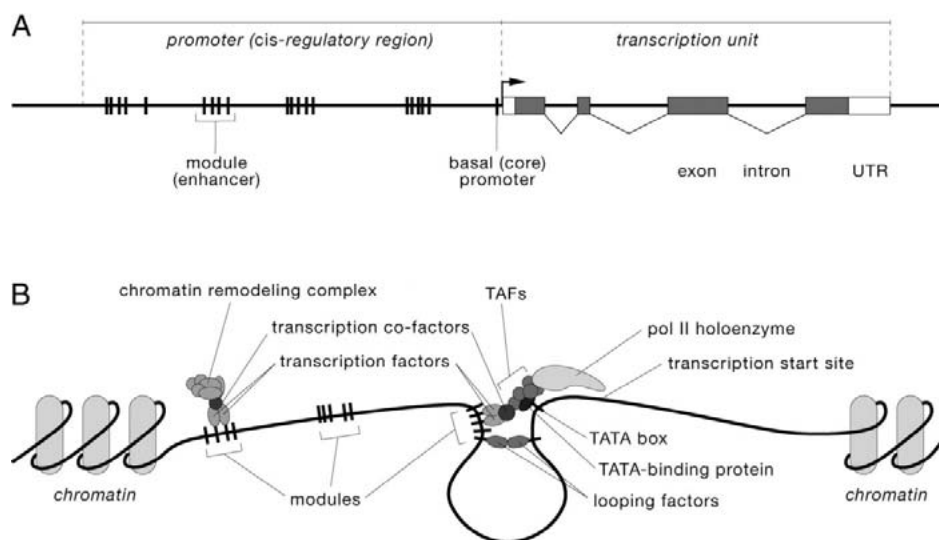


Figure 2. Promoter structure.

Top panel (A) shows a gene (transcription unit) with its 5' proximal promoter. Binding sites in the DNA for transcription factors are indicated as vertical bars. The transcription start site is indicated with the bend arrow. In the bottom panel (B) is depicted how transcription factors bind to the DNA and interact. DNA is made accessible through the actions of chromatin remodeling proteins, other transcription activating factors (TAFs) form complexes that lead to the activation of the RNA-polymerase II.

Figure adapted from: Wray et al.[5] © 2003 by the Society for Molecular Biology and Evolution. *MBE* Online: <http://www.mbe.oxfordjournals.org/>

Identifying regulatory elements

At the onset of our studies, no prior studies had been undertaken to investigate the transcriptional regulation of cardiac-specific genes as a group. Only for a few it was known which elements within their promoter brought about their individual tissue-specific expression profile.[6-10] Instead of focusing on only a single gene, we set out to investigate whether the cardiac-specific genes as a group shared a

common set of cis-regulatory elements, that eventually led to restricting transcription (and thus the final expression) of genes to the heart.

The identification of individual functionally relevant regulatory elements within a single promoter can be a daunting task. A separate laboratory experiment would be required to determine the role of each and every regulatory element. Since such individualistic screening is not feasible for large scale studies, other approaches are required. Genes expressed at the same time and place are prone to recruit a shared set of TFs. Their co-expression is thus also likely to result from the presence of similar regulatory elements within their promoters. In addition, it has also been found that the regulatory mechanisms of transcription of many genes, including cardiac genes, are often conserved for the orthologues genes across species.[8, 10, 11] Based on this principle, numerous bioinformatics programs have been developed to identify which regulatory elements are enriched within the promoters of a set of co-expressed genes when compared to a background set.[12] In this way lists can be generated of motifs and or modules most likely to be involved in the transcriptional response of interest. The labor-intensive and more costly laboratory studies can then be specifically targeted to the validation of the computer derived (in silico) results (see figure 3).

Models for research of cardiac transcription

Cardiac research benefits from the fact that the heart or a similar organ is present in all higher organisms. From *Drosophila* to vertebrates there is a strong conservation in the mechanisms driving cardiogenesis.[13, 14] Since there is no 'simple' human model system to study transcription regulation in ventricular cardiomyocytes, studies in this field are typically performed using primary cell cultures of rat or mice cardiomyocytes. Like humans, also *muridae* have hearts consisting of separate atria and ventricles. As with the use of any model there are some restraints in the translation of the results from the model to for example the human situation. Despite the many similarities also some key differences exist. One of the most striking ones besides size is probably the rate of contraction. Where the number of beats per minute for the human heart is around 70, the heart rate of rat or mice can reach over 500.

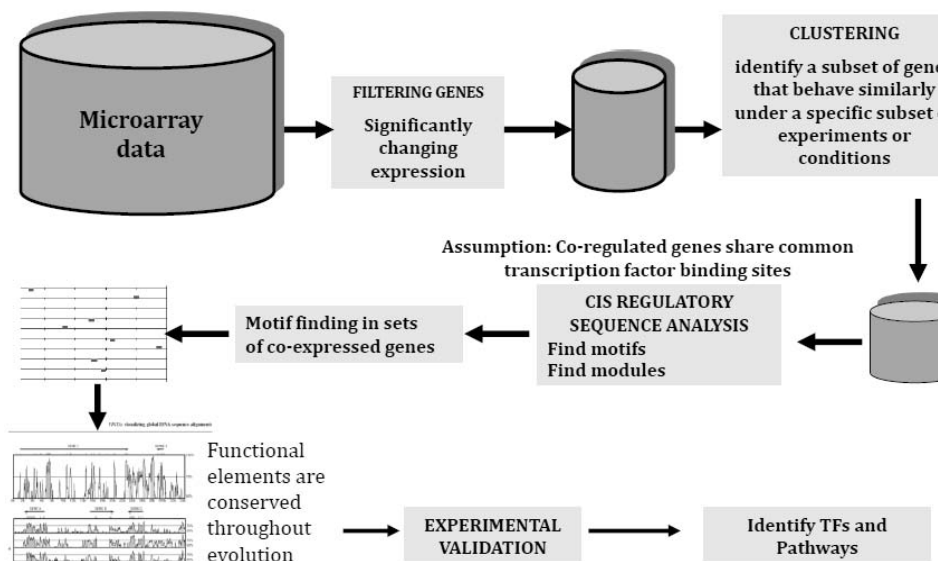


Figure 3. Flow diagram for the identification and validation of regulatory promoter motifs.

Genes expressed at the same time and place are prone to recruit a shared set of TFs. Their co-expression is thus also likely to result from the presence of similar regulatory elements within their promoters. Identifying such shared promoter elements can assist in identifying which regulatory mechanisms are involved in establishing a specific gene expression-profile under a given set of conditions.

In the absence of electrical pulsing, the functionality of cardiomyocytes in culture rapidly diminishes.[15] Applying electrical pulsing has been found to conserve the contractile activity by triggering calcium sparks.[16, 17] Additional studies showed that by applying electrical stimulation, the expression of several cardiac-specific genes is increased and the ultra-structures of the cardiomyocytes are maintained while these would be lost otherwise.[18, 19] Despite the apparent relevance of electric signals for cardiomyocytes, no study has been published that describes the effects of electric pulsing on the whole cardiomyocyte transcriptome. Furthermore, there also remains the question of to what extent the changes in transcription observed for the individual genes were the result of the electric stimuli or their consequential contractions. To fully understand the role of electrical

signaling it is therefore important to establish whether genes respond differently to electrical stimuli and contractions.

Predicting gene function using transcriptomics data

Alongside gene expression data on targets of interest, experiments with expression micro-arrays also produce data on the thousands of other genes, including some that have remained poorly annotated. These transcriptomics data can also be used to predict the functional role of genes lacking functional annotation. This is done based on the hypothesis that genes which are functionally related are subject to shared transcriptional regulatory mechanisms. For example, Walker and colleagues identified eight 'new' genes involved in prostate cancer based on their profiles of co-expression with known 'prostate cancer related genes.[20] Expression profile analyses can thus provide valuable new insights in the possible functional role of genes.

Aims of the study

Many questions still remained on various aspects of the cardiac transcriptome. In our studies we addressed the following. First, to what level does the cardiac muscle-specific gene expression program overlap with the programs of the other muscle tissues, and to what extent is it unique to the heart? Is there a common transcription regulatory program that lays at the basis of restricting the expression of genes to the cardiac muscle? Second, what is the effect of electrical stimuli at the level of the transcriptome? Which genes and associated processes are hereby affected? Third, is there, analogous to the stretch specific program, also an electrical-pulsing specific response and if so, which factors are involved?[21]

Transcriptomics data can also be used to predict gene functions. Several genes with a heart-enriched expression profile are still awaiting their functional annotation. This brings about the fourth question; can we utilize the often ample available expression data to identify the functional roles of these genes?

Besides investigating these questions we also aimed at the identification of promoters of genes for which these had not been identified yet. By determining the promoters of individual genes active in one or more of the various types of muscle, we generate the tools that

can be used in experiments to validate the findings from the genome-wide studies.

Outline of the dissertation

In **chapter 2** an overview is presented of the transcription factors involved in the activation of cardiac-specific genes and of the regulatory networks in which they exert their function.

In our studies we made extensive use of bioinformatics, however, some sort of experimental validation is eventually required to validate the results obtained. Furthermore, the set of cardiac genes for which the promoters had been described was fairly limited. Therefore we created a number of luciferase-reporter constructs based on human promoter sequences (**chapter 3**). These constructs enabled us to test tissue specificity of promoter constructs and the relevance of motifs identified in the bioinformatics analyses. Where other studies often only reported which promoter fragments of cardiac genes were sufficient for driving expression in cardiomyocytes, we investigated in addition also the degree of tissue specificity.

To investigate the effects of electrical pulsing on cardiomyocyte transcription we developed a system in which we were able to culture cardiomyocytes for longer periods while continuously subjecting them to electrical stimulation. This system and the associating culture protocol were described in **chapter 4**.

In order to identify the effects of electrical pulsing on the cardiomyocyte transcriptome on a genome-wide scale, we performed a comprehensive microarray study, presented in **chapter 5**. Besides identifying which genes are differentially expressed as a result of electric pulsing, we also distinguished in our study whether differential expression was resulting from the electrical pulsing on itself apart from the evoked contractions.

In **chapter 6** we described how bioinformatics analyses were employed to identify which regulatory elements are responsible for restricting the expression of genes to the heart. In addition, it presents potential candidate motifs responsible for the eventually link up of the electrical signals with the transcription machinery.

TCF7L2 which was identified in chapter 6 as a modulator of the transcriptional response to electrical pulsing, is a well known target of β -catenin signaling.[22] In **chapter 7**, we reported on how we further investigated the relation of this signaling pathway with that activated by

electrical pulsing, and also on the possibility to manipulate these mechanisms by using pharmacological compounds.

As an example of how gene-expression data can be employed to identify the functional role of a gene, we used the approach of guilt by association to elucidate the function of a mitochondrial gene with a heart-enriched expression profile (**chapter 8**).

Finally, **Chapter 9** provides an overall discussion of the research presented in this dissertation.

References

1. A. West, P. Shore, and A. Sharrocks, DNA binding by MADS-box transcription factors, a molecular mechanism for differential DNA bending. *Mol Cell Biol.*, 17 (1997) 2876-2887.
2. J. Davie, Nuclear matrix, dynamic histone acetylation and transcriptionally active chromatin. *Molecular Biology Reports*, 24 (1997) 197-207.
3. R.G. Roeder, The role of general initiation factors in transcription by RNA polymerase II. *Trends in Biochemical Sciences*, 21 (1996) 327-335.
4. R.G. Roeder, Transcriptional regulation and the role of diverse coactivators in animal cells. *FEBS Letters*, 579 (2005) 909-915.
5. G.A. Wray, M.W. Hahn, E. Abouheif, et al., The Evolution of Transcriptional Regulation in Eukaryotes. *Mol Biol Evol.*, 20 (2003) 1377-1419.
6. J.D. Molkentin, S.M. Jobe, and B.E. Markham, alpha-myosin Heavy Chain Gene Regulation: Delineation and Characterization of the Cardiac muscle-specific Enhancer and Muscle-specific Promoter. *J Mol Cell Cardiol.*, 28 (1996) 1211-1225.
7. M.N. Daniel Durocher, Combinatorial interactions regulating cardiac transcription. *Developmental Genetics*, 22 (1998) 250-262.
8. A. Minty and L. Kedes, Upstream regions of the human cardiac actin gene that modulate its transcription in muscle cells: presence of an evolutionarily conserved repeated motif. *Mol Cell Biol*, 6 (1986) 2125-36.
9. M. Lemonnier and M.E. Buckingham, Characterization of a cardiac-specific enhancer, which directs alpha-cardiac actin gene transcription in the mouse adult heart. *J Biol Chem.*, 279 (2004) 55651-8.
10. H. Zhu, A.V. Garcia, R.S. Ross, et al., A conserved 28-base-pair element (HF-1) in the rat cardiac myosin light-chain-2 gene confers cardiac-specific and alpha-adrenergic-inducible expression in cultured neonatal rat myocardial cells. *Mol. Cell. Biol.*, 11 (1991) 2273-2281.
11. W.W. Wasserman and J.W. Fickett, Identification of regulatory regions which confer muscle-specific gene expression. *Journal of Molecular Biology*, 278 (1998) 167-181.
12. S. Aerts, G. Thijs, B. Coessens, et al., Toucan: deciphering the cis-regulatory logic of coregulated genes. *Nucleic Acids Res*, 31 (2003) 1753-64.

13. E.N. Olson, Gene regulatory networks in the evolution and development of the heart. *Science*, 313 (2006) 1922-7.
14. R.M. Cripps and E.N. Olson, Control of Cardiac Development by an Evolutionarily Conserved Transcriptional Network. *Dev Biol.*, 246 (2002) 14-28.
15. H.J. Berger, S.K. Prasad, A.J. Davidoff, et al., Continual electric field stimulation preserves contractile function of adult ventricular myocytes in primary culture. *Am J Physiol.*, 266 (1994) H341-349.
16. H. Cheng, W.J. Lederer, and M.B. Cannell, Calcium sparks: elementary events underlying excitation-contraction coupling in heart muscle. *Science*, 262 (1993) 740-4.
17. E. Holt, P.K. Lunde, O.M. Sejersted, et al., Electrical stimulation of adult rat cardiomyocytes in culture improves contractile properties and is associated with altered calcium handling. *Basic Res Cardiol.*, 92 (1997) 289-298.
18. P.M. McDonough and C.C. Glembotski, Induction of atrial natriuretic factor and myosin light chain-2 gene expression in cultured ventricular myocytes by electrical stimulation of contraction. *J Biol Chem.*, 267 (1992) 11665-11668.
19. M. Radisic, H. Park, H. Shing, et al., Functional assembly of engineered myocardium by electrical stimulation of cardiac myocytes cultured on scaffolds. *PNAS*, 101 (2004) 18129-18134.
20. M.G. Walker, W. Volkmuth, E. Sprinzak, et al., Prediction of Gene Function by Genome-Scale Expression Analysis: Prostate Cancer-Associated Genes. *Genome Res.*, 9 (1999) 1198-1203.
21. D. Frank, C. Kuhn, B. Brors, et al., Gene expression pattern in biomechanically stretched cardiomyocytes: evidence for a stretch-specific gene program. *Hypertension*, 51 (2008) 309-18.
22. J. Behrens, J.P. von Kries, M. Kuhl, et al., Functional interaction of [beta]-catenin with the transcription factor LEF-1. *Nature*, 382 (1996) 638-642.

CHAPTER 2

REGULATORY NETWORK OF THE CARDIOMYOCYTE TRANSCRIPTOME

Ruben S.R.M. Martherus, Torik A.Y. Ayoubi

Submitted

Abstract

Constituting one of the three muscle types in the human body, cardiomyocytes share many similarities with the other muscles but at the same time display several unique features. The unique properties are obtained through the activity of several genes of which the expression is highly enriched in or even completely limited to only the cardiomyocytes. At the basis of the tissue-specific expression of these genes lays an intricate network of transcription factors which in part is shared with the skeletal and smooth muscle cells and in addition involves several key factors that to a large extent are restricted to cardiomyocytes. While most overviews are primarily focused on the actions of individual transcription factors here, we deliberate on interactions amongst the collective of transcriptional initiators, inducers and inhibitors involved in regulating the cardiac expression at the transcription level, taking into account the different levels of tissue specificity and the possible modes of action of the factors involved. The picture that emerges is that of a core network with MEF2 at its center. Together with key cardiac-enriched players like GATA4/6, NKX2-5, HAND2 and more ubiquitous factors like SRF and additional factors it controls the regulation of many of the heart-specific genes. Additional activators and a variety of inhibitors provide further fine tuning of transcription and together form the eventual regulatory machinery that gives rise to the formation of the complex organ that is the heart.

Introduction

The human body contains three different highly specialized types of muscle: cardiac, skeletal and smooth muscle performing their function at different locations, controlled by distinctive mechanisms. Smooth muscle enables the contraction of blood vessels, the gastro-intestinal tract and organs like the bladder and uterus. Its relatively slow contractions are controlled by the autonomous nervous system in contrast to the fast contractions of skeletal muscles that are controlled by the somatic nervous system and the continuous rhythmic contractions of the heart muscle initiated upon signals originating in the sinus node. The differences in contractile characteristics require adaptations in the architecture of the contractile apparatus, metabolism and other elements involved such as ion channels within and in between the cells. Thus apart from all being types of muscle and consequentially

sharing many similarities, the gene expression programs of the muscles also have very distinctive features.

In this review we explore the regulatory mechanisms at the transcription level that lead to the restriction of gene expression within cardiomyocytes, the heart's unique muscle cells. Only for a small number of genes expressed in cardiomyocytes is known which regulatory sequences confer the restriction of their expression to this specific cell type. Numerous articles elaborate on specific transcription factors (TFs) and their role in a particular promoter of a cardiac gene. Concise overviews, however, of the bigger regulatory networks in which these factors are involved are very limited, especially the mechanisms of how these networks of interacting proteins lead to the tissue-specific transcription activation. Here we provide an integral overview of the transcription regulatory networks leading to cardiomyocyte specific expression of various genes.

From general to cardiac-specific transcription

Regulating transcription is one of the key control mechanisms in the process of gene expression. Most protein encoding genes depend for their transcription on the activation of RNA polymerase II (RNAPolII).[1] In association with several fairly ubiquitous expressed factors, RNAPolII is capable of driving transcription at low (basal) levels in cell-free systems.[2] Because of some functional redundancy there exist quite some variation amongst promoters in which of these core-factors are employed.[3] Some of these factors are required to bind to specific DNA sequences within the DNA, the so-called core-promoter elements such as: TATA-box, Initiator (Inr) and downstream promoter element (DPE).[4, 5] To obtain expression levels above that of the basal expression additional, often more gene-specific TFs, are required.[6] These gene-specific TFs are shared by multiple genes with a similar promoter structure. As a result the DNA of the activated genes is relocated to specific sites within the nucleus where they become transcribed within close proximity of one another: the transcription factories.[7] Also active RNAPolII readily resides at high concentrations in a hyper-phosphorylated form at transcription factories and the factories are maintained even in absence of transcription.[8, 9]

Just like many other proteins also the expression of TF-encoding genes, especially the gene-specific, occurs in a spatial and temporal specific manner. Enrichment of gene expression in a tissue can be

scored through calculating entropy whereby the expression in one tissue is compared to that in a set of other tissues. To determine the enrichment of all transcription factors within the heart we determined their Shannon entropy values as described by Schug et al.[10] For the calculations we made use of an extensive publicly accessible dataset from the Neurocrine corporation that contains expression data over 65 different human tissues.[11] All TFs were selected based on being annotated with the gene ontology (GO) term of 'transcription factor activity'. The entropy for genes in heart ventricle and atria were determined separately since although both being an primary component of the heart, the structures are very different from one another and some transcription factors display a strong specificity for one of the two.

In table 1 we present the twenty most ventricle-specific transcription factors and in addition indicated their ranking in atria. It can be seen clearly that several factors such as GATA4 and GATA6 exhibit somewhat similar tissue specificity for ventricle and atria while others, HEY2 (hairy/enhancer-of-split-related with YRPW motif 2) in particular, seem to be highly specific for ventricle. Furthermore it becomes apparent that some isoforms encoded by the same gene, e.g. TBX5 (T-box 5), display different degrees of tissue specificity for ventricle and atria. In figure 1 we present a graphical display of a selection of TFs present within 1 kb of the proximal promoter of cardiac and non-cardiac genes. In addition, we calculated for each gene the Shannon entropy values for cardiac ventricle, skeletal muscle and the cerebellum. Indeed the proximal promoters of the cardiac genes are enriched with binding sites for TFs like NKX2-5 and GATA. However, more noticeable is the abundance of SP1, SRF and MEF2. Although none of them ranks amongst the top cardiac-specific TFs, they do play a pivotal role in determining cardiac specificity by recruiting more cardiac-specific (co-) factors. Studies on the mechanisms of regulating cardiac genes were often focused on either *NPPA*, encoding the precursor peptide of atrial natriuretic factor (ANF) or cardiac alpha actin (*ACTC1*).[12, 13] In the following sections we elucidate on their regulation and the interactions of the TFs involved. A graphical representation of all mentioned interactions between TFs is given in figure 2.

Table 1. Ranking of transcription factors according to tissue specificity in adult heart ventricle and atrium

Rank ventricle	Rank atrium	HGNC ID	Gene Title
1	1	Gata6	GATA binding protein 6
2	4	Gata4	GATA binding protein 4
3	7	Nkx2-5	NK2 transcription factor related, locus 5 (Drosophila)
4	6	MYOCD	myocardin
5	1071	Hey2	hairy/enhancer-of-split-related with YRPW motif 2
6	38	Sox7	SRY (sex determining region Y)-box 7
7	2	TBX5	T-box 5
8	51	UHRF1	ubiquitin-like with PHD and ring finger domains 1
9	26	CSDA	cold shock domain protein A
10	22	Mitf	microphthalmia-associated transcription factor
11	78	Sox7	SRY (sex determining region Y)-box 7
12	63	Mitf	microphthalmia-associated transcription factor
13	113	Hand1	heart and neural crest derivatives expressed 1
14	30	Heyl	hairy/enhancer-of-split related with YRPW motif-1
15	73	ace2	angiotensin I converting enzyme
16	36	Epas1	endothelial PAS domain protein 1
17	53	Nr4a2	nuclear receptor subfamily 4, group A, member 2
18	5	TBX5	T-box 5
19	18	MSRB2	methionine sulfoxide reductase B2
20	27	Mef2a	Myocyte enhancer factor 2a

Genes are ranked by increasing Shanon entropy values calculated from average gene-expression in the dataset of Roth et al.(2006) encompassing 353 samples from 65 tissues.

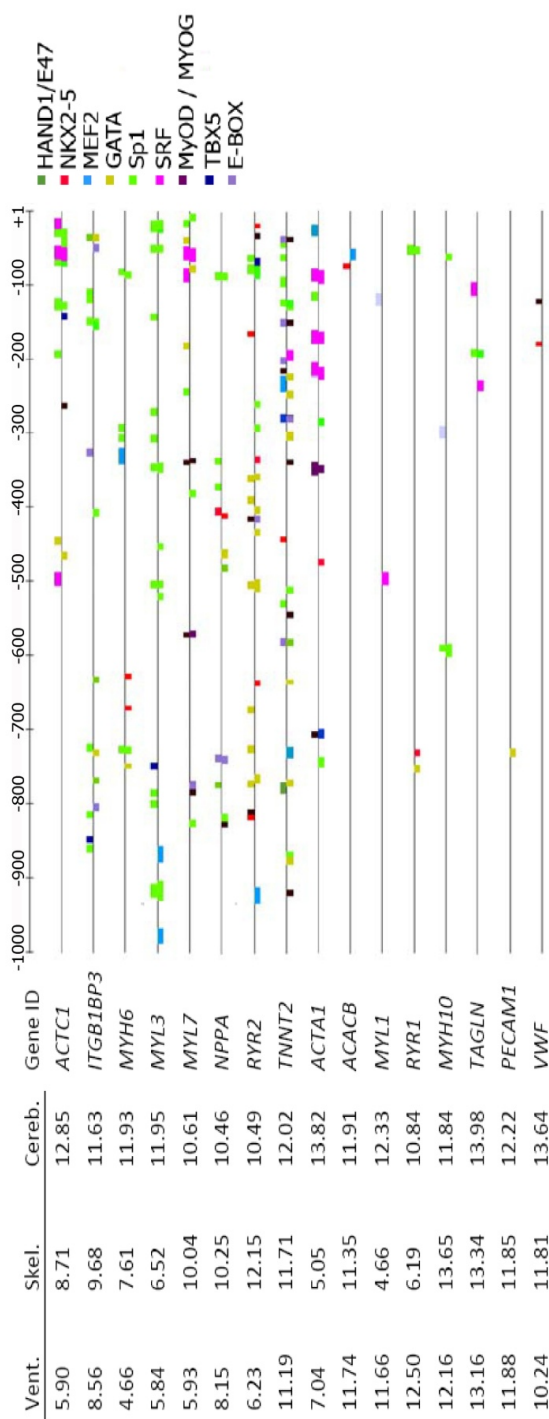


Figure 1. Enrichment of regulatory elements in promoters of cardiac genes.

Of all indicated genes 1 kb of the directly 5' proximal genomic DNA sequence was retrieved and scanned for the presence of potential transcription factor binding sites of HAND1, NKX2-5, MEF2, GATA4/6, Sp1, SRF, MyoD/MYOG, TBX5 and E-box binding factors. Genes and tissues in which they are enriched: ACTC1 to TNNT2 - heart muscle, ACTA1 - striated muscle (heart and skeletal), ACACB to RYR1 - skeletal muscle, MYH10 and TAGLN - smooth muscle and PECAM1 and VWF - endothelium. Table displays Shannon entropy values in cardiac ventricle (Vent.) skeletal muscle (Skel.) and Cerebellum (Cereb.) calculated from average gene-expression in the dataset of Roth et al.(2006) encompassing 353 samples from 65 tissues.

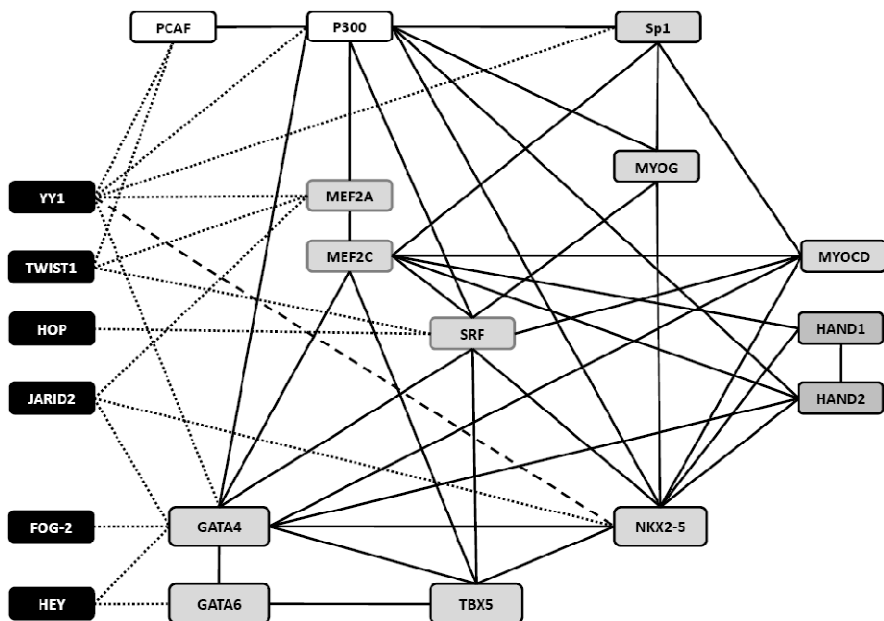


Figure 2. Overview of regulatory networks between transcription factors. Each line represents direct interaction between two transcription factors, additional factors that may be required for the interactions are not indicated. Associations given as continuous lines lead to activation of target genes, dotted lines indicate repressor activity by factor with inhibitory function (black boxes). Nkx2-5 can replace YY1 at target sequences, thereby removing YY1's inhibitory activity (dashed line). Names of factors highly enriched in cardiomyocytes are underlined.

Regulation of the *NPPA* promoter

Although the name might suggest otherwise, *NPPA* expression is not restricted to the atria, but also occurs in the heart ventricles, lungs, liver and pituitary gland.[14] The proximal promoter fragment of 700 bp appears to be sufficient for driving cardiac expression of *NPPA* (see figure 3).[15]

NKX2-5

In the *NPPA* promoter a key role is played by the cardiac-enriched transcription factor NKX2-5 the vertebrate homologue of the *Drosophila* Tinman (*Tin*) which is crucial for cardiogenesis.[16] Its expression precedes that of the cardiac-specific genes alpha myosin heavy chain (*MYH6*), *NPPA* and *ACTC1*. [16-18] In vertebrates *NKX2-5* is expressed during all the stages of life, initially throughout different tissues but in later stages of life it becomes more restricted to the heart.[17, 19] NKX2-5 is a member of the NK2 class of homeodomain/ homeobox (Hox)-containing transcription factors. Hox-proteins on themselves often do not have high binding affinity but when bound by co-factors binding specificity and affinity are increased.[20] NKX2-5 can transactivate the mouse *Nppa* promoter through binding to two separate binding sites on the DNA, named NKX2-5 response elements (NKEs) with sequences CCAAGTG and TCAAGTG located at -91/ -78 and in tandem at -243/-221 bp from the TSS.[20, 21] NKX2-5 binds to different sites on the DNA with varying affinity, binding is strong to TNNAGTG and weaker to CWTTAATTN.[22] Together with the zinc-finger protein GATA4, NKX2-5 is capable of activating the *NPPA* promoter in non-cardiac cells.[23] This and the ability of NKX2-5 to transactivate the *Nppa* promoter in fibroblasts demonstrate a crucial role of NKX2-5 in the activation of the cardiac gene expression program. Higher affinity binding of NKX2-5 to the DNA can be obtained via interaction with GATA4: this protein induces a conformational change in that protein that takes away the inhibitory effect of the NKX2-5 C-terminal domain so that p300 can interact with the NKX2-5 N-terminal domain thereby contributing to DNA binding affinity.[24]

GATA4 and p300

Like NKX2-5, GATA4 is expressed in only a few tissues. In adults, expression is only found in the heart, liver, small intestine, ovary and testes.[15, 20, 25] GATA factors contain one or two zinc-finger domains that together with an adjacent conserved highly basic region form a domain that binds DNA to sequences containing the GATA motif with the consensus sequence WGATAR. The importance of GATA4 for *NPPA* transcription becomes apparent from the fact that *Nppa* promoter activity in mesodermal 10T1/2 cells requires direct binding of GATA4 to the DNA. Out of the two binding sites that are present: at -138 and -295 bp, the proximal one is the most essential.[20, 26] GATA4 can

functionally be replaced to some extent by GATA6.[27] The two proteins also directly interact with each other thereby enhancing *NPPA* transcription, requiring only one of the factors to bind their shared consensus motif on the DNA.[28] Besides collaborating with other GATA factors, the GATA proteins act in synergy with a wide variety of other proteins besides NKX2-5. DNA bound GATA4 can recruit MEF2 to target promoters without requiring MEF2 itself to bind to DNA.[29] Together with the serum response factor (SRF) GATA4 can activate various SRE (serum response element, binding site for SRF) dependent promoters.[30]

It has been shown that this GATA dependent activation of transcription involves recruitment of the general TF p300 that has histone acetyltransferase (HAT) activity and interacts with the basal transcription initiation complex (TIC) of RNA-polymerase II through binding to TBP.[31, 32] GATA4 is also likely to be involved in enabling binding of p300 to NKX2-5. By removing the impeding effect of NKX2-5's C-terminal domain GATA4 might enable direct interaction between NKX2-5 and p300.[24, 33] Other regulators of GATA factor activity are the serum response factor (SRF), NFAT and Friend of GATA (FOG) proteins.[31, 34, 35] The long isoform of the FOG-2 directly inhibits *NPPA* promoter activation by GATA4.[36] While this inhibitory long isoform occurs in many types of tissues the shorter more heart restricted isoform has no inhibitory effect on *NPPA* transcription. Besides *NPPA*, FOG-2 also influences transcription of brain type natriuretic peptide (*BNP*) and the heart specific SERCA-channel encoding gene *ATP2A2*. [36, 37]

SP1

In a rat hypertrophy model, it was found that activation of the *Nppa* promoter by GATA4 could be further increased through synergetic action with Sp1, a ubiquitous expressed zinc-finger protein.[38] For this activation SP1 must bind directly to SP1 binding sites within the proximal promoter. SP1 was shown to interact with GATA4 through direct protein-protein binding. Furthermore it was shown that recruitment of SP1 to its binding sites was increased at higher levels of GATA4.[38]

MEF2

The expression of several cardiac proteins in mice, including *NPPA*, requires the MADS-box protein myocytes enhancer factor 2C (MEF2C).[39] The MADS box domain facilitates binding to DNA and interaction with other TFs including dimerization with other MADS-box proteins. The distinguishing feature of the MEF2 proteins is the MEF2 domain, a 29-amino acids domain that enables MEF2 interaction with other accessory factors and binding to its own specific DNA binding site.[40-43] MEF2 does not dimerize with other MADS-box factors such as SRF, but only forms homo and heterodimers with MEF2 proteins.[44] Similar to other MADS-box factors, MEF2 DNA-binding and activity is greatly influenced by sequences adjacent to the TFBS, in muscle-specific enhancers the overall MEF2 consensus sequence YTTA(W)₄TAR has been found to be regularly flanked by a G/C rich motif termed the CCAAC-box.[45] The DNA binding specificity of the factors themselves is determined by the first amino acid of their MADS-box domain.[44, 46, 47] However, there is some competition for the MEF2 site between MEF2 and ubiquitous nuclear factors.[48]

There is a big variety of possible combinations of MEF2 proteins due to fact that many MEF2 isoforms are generated by alternative splicing in the transcription mediating C-terminal domain.[44] In vertebrates four MEF2 genes have been found: *MEF2A*, *-B*, *-C* and *-D*. *MEF2A*, *-C* and *-D* have the same binding specificity although MEF2B binds the same sequence with a lower affinity compared to the others. These isoforms are to some degree tissue-specific, but all four MEF2 proteins are mainly expressed in myogenic cells and further muscle restricted transcriptional activity of MEF2 is obtained through post-translational regulation of MEF2.[49] MEF2 factors are found to be essential co-regulators of myogenic basic helix-loop-helix (bHLH) proteins.[50] In conjunction they can initiate myogenesis. In mice MEF2C was shown to be required for the expression of a number of cardiac specific genes *Nppa*, *Actc1*, *Myh6* and *Hand2*. [39, 51] MEF2 proteins can synergistically activate transcription by direct DNA binding or as co-factors.

-600	-590	-580	-570	-560	-550	-540	-530	-520	-510	
ATAGTCTTT	CACCTGACTG	CTAACAGGGA	CACTTAGGGT	GGGGTGGG	TGCTGGGG	CAGAGTTCCA	CCACAGAGGC	CGATGAATCA	GGTGTGAAGC	TBE
-500	-490	-480	-470	-460	-450	-440	-430	-420	-410	
							SP1			
TAGTCCAGC	ATGTGTACTC	CCTGGCCAGC	CTAGCTGGCC	TCCAGTGC	CTGTCAATGC	CTCTCTCCC	GCCTTTATT	GGAGCCCTG	ACAGCTGAGC	
-400	-390	-380	-370	-360	-350	-340	-330	-320	-310	
AGCAAGCAGG	GGGAGCTGGG	TGCAGGCCAG	CGTCACCCCT	CTGCTTCCCT	GCAATGGTCC	TGTTGCCAGG	GAGAAAGAAT	CCTGAGGCGA	CGCCCCAGGA	
-300	-290	-280	-270	-260	-250	-240	-230	-220	-210	
			TBE	NKE					SP1	
AGATAACCAA	GGACTCTTTT	CTGCTCTCT	CACACCTTTG	AAGTGGGGC	CTCTGAGGC	AAATCAGCAA	GAATGTGACT	CTTGCAGCTG	AGGTCTGGG	
-200	-190	-180	-170	-160	-150	-140	-130	-120	-110	
	SP1			TBE		GATA				
GGAGGGGGT	GAGTGGAGCT	GCTCAAGGCA	AAGGGGCCGT	GACAAGCTTT	GCCGAAGTGA	TAACTTTAAA	AGGCACTCTT	CTGCTGGCTC	CTCACTCCAT	
-100	-90	-80	-70	-60	-50	-40	-30	-20	-10	
	NKE		SP1				TATA-box			
CGCTTTATCG	CTGCAAGTGA	CAGAAATGGG	AGGGTTCTAG	CCCCCCTGCC	TTCTCAAAGA	GCTGGGGGGC	TATAAAAACG	GGAGATGCTG	GCAGCTAGGA	
10	20	30	40	50	60	70	80	90	100	
GACAGTGACG	GACAAAAGCT	GGAGAGAGAG	GAGAAAGAAA	CCAGAGTGGG	CAGAGACAGC	AAACATCAGA	TCGTGCCCCG	ACCCACGCCA	GCATGGGGCTC	

Figure 3. Mouse *Nppa* promoter.

Genomic DNA sequence -600 to +100 bp relative to the transcription start site. Indicated are the binding sites for TBX5 (T-box elements, TBE), GATA-factors (GATA), NKX2-5 (NKE) and SP1 (SP1). In addition to the TFs for which binding sites are indicated, *Nppa* is known to be controlled through the actions of MEF2 and HAND2.

For full activity MEF2C requires p300, the exact way of how this activation is obtained remains to be elucidated.[52, 53] In addition to bHLH proteins, it has also been observed that Sp1 and the HAND factors bind to MEF2. A combination of MEF2C and HAND2 is able to drive expression of *NPPA* in HELA cells, demonstrating the central role of MEF2 in *NPPA* expression.[42, 54] Another TF of the basal machinery targeted by MEF2 is the transcription elongation factor p-ETFB, which suggests that MEF2 can also affect the transcription after the initial initiation phase.[55]

HAND

HAND1 (Thing1, eHAND, HXT) and HAND2 (dHAND, HED, Thing2) are bHLH factors involved in determining the right and left ventricle formation.[56] The name HAND is derived from the tissues where the protein is expressed in embryonic stage: *heart*, *autonomic nervous system* and *neural crest derivatives*. [57] HAND1 is ventricle specific while HAND2 is expressed in both ventricles and atria of the human heart.[57-60] HAND1 and HAND2 can form homo- and hetero-dimers with themselves and other class B bHLH factors. Thattaliyath et al. found that HAND2 transcriptional activity can be independent of DNA-binding or presence of an E-box, but likely occurs in a hetero-dimer with an E-protein.[61] E-boxes are frequently occurring motifs with the consensus sequence CANNTG targeted by a wide variety of transcription factors.[62] In addition to this they showed that HAND2 synergistically activates *NPPA* transcription together with NKX2-5 but not with GATA4. The experiments however, were performed with a -258 fragment of the *NPPA* promoter and by using HEK 293 cells. In similar experiments, in neonatal rat cardiomyocytes and with the use of a -623 promoter fragment Dai et al. determined that in the presence of p300, HAND2 and GATA4 do interact synergistically. Just as for the interaction of HAND2 with NKX2-5 the synergistic activity was independent of an E-box but depended on the GATA-site.[63] Just like GATA4 also HAND2 binds directly to p300.[34] Both HAND1 and HAND2 can activate the *NPPA* promoter in conjunction with MEF2.[54, 64] For transactivation by HAND1, only MEF2 binding sites in the DNA are required. Binding of HAND1 did not increase DNA binding of MEF2, therefore it is believed that the increase in promoter activity is obtained through additional factors that are being attracted to the promoter by HAND1.[64]

TBX5

Another highly cardiac-specific factor activating the *NPPA* promoter is TBX-5, a member of the Brachyury/T-box family of transcription factors. The T in the name was derived from the short tail phenotype of mice that were heterozygous for one of the T-box family members and the T-box itself is a sequence specific DNA-binding domain.[65] In mammals 18 T-box factor genes have been identified, six of them (Tbx1/2/3/5/19/20) are active in the developing heart.[66] TBX proteins are essential in establishing the cardiac architecture by contributing to the formation of the separate ventricles.[67] While TBX20 is expressed in the right ventricle, Tbx5 is expressed in the developing left ventricle and was shown to be involved in the formation of the septum between the two ventricles.[67] Knockdown experiments in zebrafish suggest that direct binding of TBX5 with MEF2C is required for TBX5 to carry out its' function in cardiogenesis. In this interaction, MEF2C binds directly to the DNA-binding domain of TBX5 without the need of an MEF2-binding site.[68]

As a result of alternative splicing two different forms of TBX5 exist with apparently different roles.[69] While the shorter TBX5a is more active in proliferating cells, the longer TBX5b form is higher expressed in differentiating cells. With progression into the adult stage the ratio TBX5b/TBX5a in the heart increases, on average the TBX5-b isoform ranks higher when scoring for cardiac specificity (see table 1).[69] Transactivation of the *NPPA* promoter is only obtained through the shorter TBX5a form and not through the longer TBX5b. Tbx5 directly binds the *NPPA* promoter via T-box binding elements (TBEs) of which there are three present in the *NPPA* promoter between positions -487 and -80 bp, with the respective binding site sequences GGTGTGA, TCACACC and GTGACA.[70] These three sites enable, *at least in vitro*, a dose/effect response whereby *NPPA* expression increases with increasing levels of TBX5. Two highly similar TBEs that were identified in the connexin-40 promoter have the sequences GTGGGA and GTGACA.[70] TBX5 acts co-operatively with the GATA4 and NKX2-5 transcription factors, both discussed earlier in this review. NKX2-5 has been shown to directly bind to TBX5 and synergistic activity of TBX5 and GATA4 is known to increase *NPPA*-promoter activity in chicken.[67, 70, 71]

Activation of the cardiac actin promoter

Another well documented cardiac-specific promoter is that of the *ACTC1* gene encoding cardiac actin, a major structural protein of the sarcomeres. While in humans the gene is only active in the heart, mice and rat also display expression of this gene in the skeletal muscle. This illustrates the overlap in some of the transcription regulatory mechanisms of heart and other muscles. For *ACTC1* the immediate proximal promoter is sufficient for directing striated muscle-specific transcription activation.[72] Within the region -110 to +68, three essential transcription factor binding sites (TFBS) were identified: the Serum Response Element (SRE), a GC-rich box and an enhancer-box (E-box). The respective binding factors for these sites are SRF, SP1 and the class B myogenic bHLH proteins MyoD and myogenin (see figure 4).[73] Without any of these three TFBS and their associated factors promoter activity is completely abolished due to a failure in the formation of the transcription initiation complex (TIC).[74]

Myogenic factors

The class B basic helix loop helix-proteins MyoD and the related myogenin (MYOG, myogenic factor 4), Myf5 and MRF4 have been found to be the main regulators for differentiation of skeletal muscle cells in both developing and adult organisms.[75, 76] These factors are also capable of converting various non-muscle cells into muscle and are therefore referred to as myogenic proteins.[77] Muscle differentiation is initiated upon arresting the cell division cycle in the G₀ phase by MyoD.[76] For this arrest, and subsequent muscle differentiation, MyoD is required to form a multiyear complex with two histone acetyltransferase (HAT) factors: the ubiquitously expressed p300 and the more muscle restricted PCAF. Interestingly, only the histone acetyltransferase (HAT) activity of PCAF and not of p300 is required in the process.[78, 79] It was proposed by Berkes and co-workers that Pbx, which is bound constitutively to inactive chromatin, marks genes for activation by MyoD. MyoD activates the myogenin promoter through interacting with Pbx/Meis1 when in a heterodimer with E12.[80]

Cardiac gene expression regulation appears only to employ myogenin and not MyoD. Early murine cardiac muscle completely lacks expression of both myogenic proteins MyoD and myogenin.[81] Myogenin however, was found to be expressed in the heart of adult *Muridae* and also as to be a component of the complex of transcription

factors that regulate *ACTC1* transcription in humans.[74, 81, 82] Most class B myogenic factors cannot form homo-dimers efficiently and require the more ubiquitously expressed class A bHLH E-proteins to form hetero-dimers and thus become active.[61, 83, 84] In order to be able to function in the TIC of the *ACTC1* promoter, myogenin has to form a hetero-dimer with E-12 so it can bind to SRF, another essential factor for *ACTC1* transcription.[85, 86]

Besides bases making up the E-box itself, also the two flanking bases are essential determinants for susceptibility to the myogenic factors and determine whether they act as activators or inhibitors. The two flanking bases in cis-orientation of the E-boxes are essential for this activity: while the sequence 5'- AGG CAGGTGGC - 3' can function as a MyoD dependent enhancer, the sequence 5' - CTGCAGGTGTT - 3' cannot be activated by MyoD.[87]

Sp1

Although Sp1 plays a central role in *ACTC1* regulation, it has little effect on *ACTC1* expression when over-expressed in fibroblasts.[23] Sp1 has been reported to interact functionally with several of the factors known to be essential for transcription in cardiac muscle: myogenin, SRF and MEF2. Both MEF2C and MEF2D can be directly bound by Sp1, leading to a synergistic activation of transcription. Sp1, SRF and myogenic bHLH together form a complex.[86, 87] Sp1 can bind the CCAC boxes which flank MEF2 sites with a muscle-specific enhancer function.[42] As with many other TFs, also the synergistic effect of Sp1 combined with a MEF2 factor can occur without direct DNA binding of MEF2. Binding of Sp1 to a Sp1-specific site is sufficient to enable MEF2 recruitment.[88]

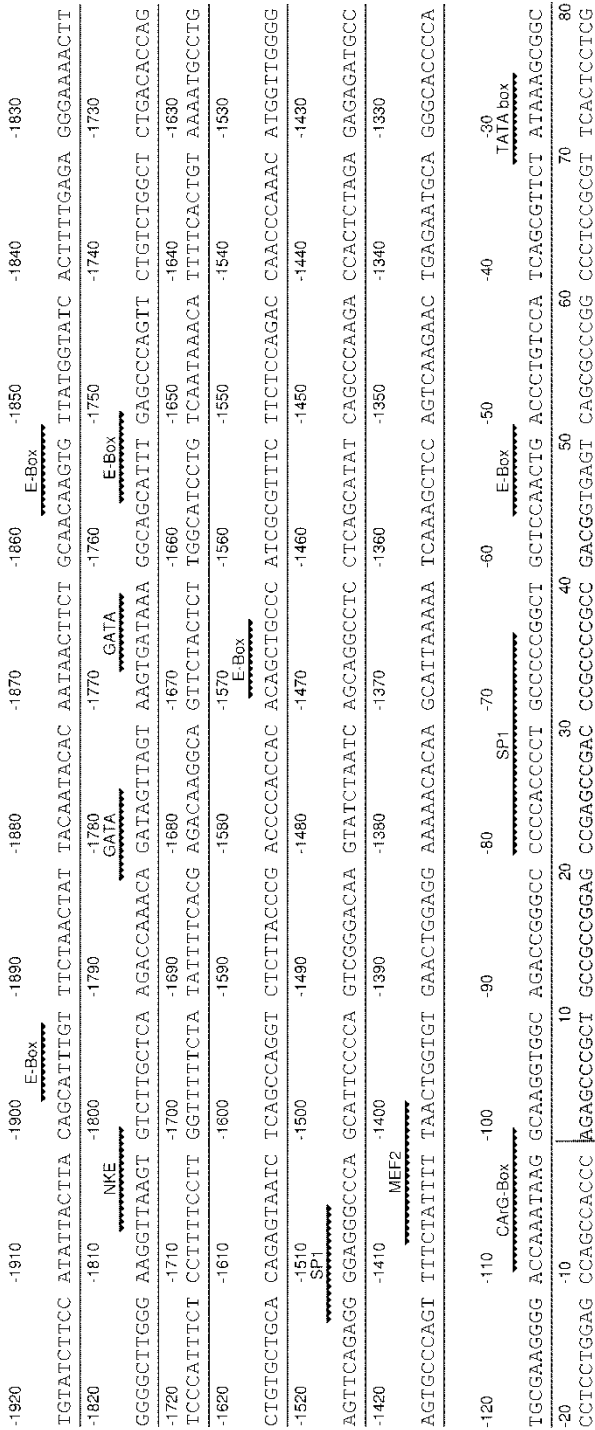


Figure 4. Human *ACTC1* promoter architecture

Two fragments of the human *ACTC1* promoter. The top, further upstream segment, is part of the region -1360 to -2353 bp that acts as a cardiac-specific enhancer. The MEF2 site at -1398 to -1407 is targeted by MEF factors. Other elements indicated have not been identified as being involved in *ACTC1* regulation although they are known target sites for transcription factors involved in regulation of cardiac genes. Bottom segment: promoter fragment of *ACTC1* sufficient for driving striated muscle-specific transcription. CA(G box, Sp1 site and E-box, bound by respectively SRF, Sp1 and myogenin.

SRF

SRF is a crucial factor in the transcription regulation in all muscle types. The protein often plays a key role, but does not determine tissue specificity.[89] Like the MEF2 factors it belongs to the MADS-box proteins family and binds to the common MADS-box binding sites with consensus sequence CCW₆GG.[47] For the activation of promoters SRF often requires the recruitment of (co)factors for full activity, even when binding to SRE sites.[90] In skeletal muscle SRF is required for the activation of genes by the two myogenic factors MyoD and myogenin thereby depending on SRE sites and E-boxes.[77] MyoD and myogenin bind SRF as hetero-dimers with E12.[86] For the activation of the *ACTC1* promoter, binding of SRF to the most proximal SRE appeared to be essential.[85] These SREs can be bound by a complex consisting of SRF and NKX2-5 as well as the two proteins independent from each other.[82, 91] However, in order to activate the promoter, SRF needs to recruit NKX2-5 without the need for NKX2-5 itself to bind to the DNA.[23, 82] The requirement for SRF to interact with other TFS suggest that a wider context of positional spacing of CArG boxes within the promoter might be essential for the formation of a multi-protein complex eventually inducing RNAPIII activity.[92]

Fine tuning ACTC1 cardiac specificity

Although SRE sites can recruit the cardiac-restricted NKX2-5, independent from- and in combination with SRF, the cardiac-specific transcription regulation of *ACTC1* requires additional regulatory elements. None of the two putative GATA binding sites have been shown to bind GATA4 and deletion of them did not result in loss of activation by GATA4.[33] For the induction of the *ACTC1* promoter, GATA4 can function as a cofactor while NKX2-5 is bound to the SREs.[20, 33]

Two DNase I hypersensitive regions upstream of the TSS of *ACTC1* function as enhancers. At about -7 kbp, a site is present for the ubiquitous factor Embigin (Emb), that is capable of recruiting MEF2 in conjunction with p300 and Oct1 which binds to an adjacent site.[93] The second enhancer, located at -2.353 to -1.36 kbp, is only active in cultured cardiomyocytes and not in differentiated skeletal muscle or non muscle cells. Within this region resides a binding site for MEF2 factors, at least MEF2A and MEF2D were shown to bind to this site.[94]

Transcription inducing co-factors

Myocardin

Amongst the more cardiac-specific genes depicted in figure 1 there is enrichment in the binding sites for SP1, SRF and MEF2. Although they themselves are not cardiac-specific factors they all enable potential recruitment of myocardin, the fourth most cardiac-specific TF (see table 1) to the promoter. Myocardin occurs in two transcript forms arising from alternative splicing. The longer, strongly heart enriched transcript (see also table 1), contains an extra miniature exon that adds a premature stop-codon, leading to a shortened form of the protein.[95] The shorter transcript encoding the larger protein is enriched in smooth muscle. Neither form of myocardin is found in skeletal muscle.[95-98] In smooth muscle myocardin in conjunction with the class II histone deacetylase (HDAC) HDAC5, was shown to assist in differentiation by blocking the skeletal muscle program through inhibition of DNA-binding by the key myogenic factor MyoD.[99] In cardiac cells myocardin plays an important role in the early stages of differentiation after its transcription is activated by Nkx2-5.[100] Myocardin also has a transcription activating function, many genes in both cardiac and smooth muscle only become activated after the expression of myocardin.[95,101-103] Myocardin binds SRF via the SRF MADS-box when it is bound to CArG boxes and this binding is not influenced by the CArG box flanking sequences.

Thyroid hormone receptor

Several genes require the thyroid hormone receptor (TR) for their full activation. Two such genes, *MYH6* and *ATP2A2*, are depicted in figure 1.[104, 105] For its transcriptional activation, TR requires the direct binding of MEF2 as cofactor.[106]

Repressors contribute to tissue specificity of transcription

So far we primarily focused on the transcriptional activators of *NPPA* and *ACTC*. However, genes attain their tissue-specific transcriptional profile through a combination of activating and repressive mechanisms. A factor with a dual role of both inducer and repressor is the ubiquitously expressed TF Yin and Yang 1 (YY1). In muscle cells it primarily has been described as an inhibitor of transcription through binding to specific YY1 sites or SREs, making them inaccessible for SRF.[107] With increased differentiation YY1 activity diminishes and the SRE becomes accessible again for SRF.[107] For full repressor activity YY1 requires acetylation by p300 and PCAF so it can bind more strongly to HDACs.[108] However, HDACs can deacetylate YY1 again, thus forming a negative feedback loop.[108] In the heart repressive activity of YY1 can be lifted through the actions of the cardiac-enriched TF Nkx2-5.[109] It has been shown that YY1 can be displaced from the *ACTC1* promoter by NKX2-5 together with SRF. This mechanism of counteracting YY1 thus only occurs in those few cell-types where YY1 expression is low and/or NKX2-5 is expressed in sufficient amounts thus contributing to the tissue specificity of *ACTC1* and other promoters such as *MYH6* which like *ACTC1* are also under control of YY1.[109, 110] Where YY1 blocks transcription through rendering the promoters inaccessible, the bHLH TF TWIST represses transcription via inactivation of key TFs by directly binding to them. Mouse TWIST can bind directly to the different bHLH myogenic factors.[111] In this way TWIST inhibits differentiation of cells into muscle lineages. Other factors of which the activity is repressed through direct binding of TWIST are MEF2 and the HATs p300 and PCAF.[112]

Also in the formation of the separate compartments of the heart, the atria and ventricles, transcriptional repressors play a key role. For example the HEY TFs that are expressed primarily in the cardiovascular system and target *GATA4* and *GATA6* expression and influence their actions through direct binding without interfering with p300 binding by *GATA4*. [113, 114] The two HEY factors are distinctively active in one of the two: Hey1 is expressed in the atria and Hey2 in the ventricles (see also table 1).[115, 116] HEY2 expression assists in the differentiation of the ventricular cardiomyocytes by repressing the atrial gene expression.[117]

A TF with a similar role to that of HEY2 is *Irx4*. This homologue to the *Drosophila* Iroquois homeobox factors is also expressed solely in the

ventricles and stimulates there the ventricle specific expression under control of NKX2-5 and HAND2 while at the same time inhibiting atrial-specific gene expression.[118, 119]

Maintaining the status quo

Transcriptional inhibitors are not only important for compartmentalization but also for and maintaining a healthy status quo the cardiac tissue. Without inhibitors transcription would go on uncontrolled since, as mentioned earlier, the transcription factories with active RNAPolII maintain present. Over-activation of MEF2 could for example lead to cardiomyopathy.[120] Class II histone deacetylases (HDACs) prevent over-activation of MEF2 through direct binding to a domain on MEF2 that is also targeted by HATs.[121, 122] Another clear example is the ventricle enriched homeodomain protein HOP, required to establish an appropriate number of cardiomyocytes by balancing proliferation and differentiation. Mutations in the gene in mice lead to ruptures in the ventricular walls of mice embryos.[123] The protein was first identified as an NKX2-5 dependent homeodomain protein antagonizing DNA binding by SRF and indeed the mice with mutated HOP had altered expression levels for many SRF target genes.[18, 123, 124] Separately, HOP does not affect expression of *NPPA*, *ACTC*, but with addition of SRF the genes are down regulated in a dose dependent way, via inhibition of SRF binding to DNA. Furthermore HOP diminishes the responsiveness of *NPPA* and the smooth muscle gene *SM22* to myocardin.[123]

The transcriptional repressor JARID2 (jumonji/JMJ) has both repressive and activating properties. JARID2 represses *NPPA* transcription through direct binding to the key TFs Nkx2-5 and GATA4.[125] A stimulatory effect of JARID2 has been observed for *MYH6*, however in another study where a much smaller promoter fragment was used (368 bp instead of 5.6 kb), JARID2 was shown to repress the activity of *MYH6* transcription.[126, 127] In the latter study JARID2 was shown to bind directly to MEF2 thereby inhibiting its binding to the thyroid receptor (TR) required for the synergistic activation of the *MYH6* promoter.[106]

Another repressor acting through inhibition of TR is the heart and brain enriched 'friend of GATA'(FOG)-2.[128] FOG-2 inhibits the stimulating effect of THR-alpha on *ATP2A2* expression by direct binding to THR-alpha.[37] FOG-2 is probably better known for its repressive

effect on the transcription of cardiac-specific genes via the inhibition of GATA4, whereby a possible mechanism is that of FOG-2 competing with GATA4 for binding of p300.[129, 130]

Direct activation amidst key cardiac transcription factors

The cardiac-specific transcription program starts with the synthesis of the transcription factors involved. The early expression in the heart of NKX2-5 depends on a GATA4 binding site at about 9.3 kb upstream in the promoter.[131] GATA-4 on itself is only capable of weak activation of the promoter; higher levels of NKX2-5 expression are obtained through recruiting GATA-4 to Smad sites by Smad1 and Smad4. A similar activity was also observed for GATA6.[132]

In mouse embryonic development, *Mef2C* is a direct transcriptional target of GATA4.[133] In *Drosophila* a cardiac-specific enhancer is found 6 kb upstream of the *Mef2d* gene. This enhancer contains two binding sites for the cardiac homeodomain protein Tinman NKX2-5 homologue, pointing to a role of Tinman in the regulation of MEF2 expression. On its turn MEF2 turns on the transcription of myogenic factors. Myogenin expression which is only initiated after the start of muscle differentiation is dependent on binding of the MADS-box protein MEF2 to a MEF2 site in its promoter. Of the myogenin promoter 184 bp upstream of the TSS confers muscle specificity and auto regulation via an E-box.[87, 134] The expressed myogenin can contribute to inducing MEF2 transcription in non-myogenic cells in absence of mitogenic factors and other myogenic proteins such as MyoD.[48] Muscle differentiation proceeds further as a result from the regulatory loops existing between the myogenic factors themselves.[80] DNA binding activity of the MEF2A isoform is increased in presence of MyoD in non muscle cells. And on its turn MEF2A also induces the expression of MyoD.[135]

Transcription of SRF in developing cardiac tissue of embryonic mice depends on binding of TBX2 and TBX5 to T-box binding sites in the 3'UTR in synergy with the HAT protein HIV-1 Tat interacting protein (HTATIP). The 3'UTR in combination with a 0.5 kb 5' flanking promoter sequence was sufficient for cardiac expression.[136]

Synopsis of the regulatory network of cardiac-specific transcription

The cardiac-specific genes depend for their transcription initiation on the activation of the same sub-processes as all other genes: chromatin remodeling, formation of the RNA polymerase holo-enzyme complex and its activation at the TSS. The limitation of cardiac specificity already starts at the level of chromatin remodeling. The recent finding of the interaction between the cardiac-enriched HAND2 and Pbx suggest that cardiac genes could be marked for transcription in a way similar to that in skeletal muscle.[137] HAND2 on its turn can recruit key factors of the cardiac transcription program like MEF2, GATA4 and NKX2-5 and also additional ubiquitously expressed co-factors such as p300 which are required for the transcription activation of genes by for example GATA4. GATA4 also functions in another chromatin remodeling mechanism active in cardiomyocytes, together with TBX5 and a cardiac-specific subunit of a chromatin remodeling complex: Baf60c.[138] Once chromatin has become de-condensed various TFs can readily enhance further stabilization of the open structure through their histone acetylase function.

MEF2 appears to be at the basis of the gene expression program in striated muscles and it does so in close co-operation with the powerful myogenic transcription factors. As if being a specialized form of skeletal muscle, the cardiomyocytes only employ a subset of these factors. While in skeletal muscle MyoD has been described as the dominating TF, in the heart only myogenin is active. In a similar fashion there is some variation in which isoforms of MEF2 are active. Although the combined action some factors can result in the expression of cardiac-specific genes even in non-muscle cells *in vitro*, the endogenous promoter of all the genes described here contain functional transcription factor binding sites that are targeted by MEF2 or contain binding sites for one or more TFs that can directly recruit MEF2. MEF2's central role also becomes apparent from figure 2: at least four factors were shown to exert a direct inhibitory effect on MEF2 activity and no less than nine TFs work in synergy with MEF2. Taking into account the fact that MEF2 was shown to be required for transcriptional activity of many cardiac-specific genes and the fact that MEF2 is a direct downstream target of NKX2-5 it seems likely that also MEF2 is involved in the observed activation of cardiac-specific genes upon expression of GATA4 and NKX2-5 in non-muscle cells.

Both GATA4 and NKX2-5 emerged from studies on cardiac gene regulation as key players. Both can act in different ways: through direct binding and as cofactors and furthermore they also function upstream of many factors. Figure 5 shows the direct activation of transcription of other essential cardiac transcription factors, especially by NKX2-5. The high degree of cardiac specificity of NKX2-5 (see table 1) combined with its functions as direct DNA binding factor, widely interacting co-factor and its activity in lifting the repressor activity of YY1 make this TF a crucial player in obtaining cardiac specificity of transcription initiation. A study performed by Ieda et al., showed that the combined activity of MEF2C, GATA4 and TBX5 was sufficient for the reprogramming of fibroblasts into cardiomyocytes.[139] Although NKX2-5 was thus not amongst these, the direct trans-activation of the latter by GATA4 would still enable it to play a role in this reprogramming. A study describing the translocation of NKX2-5 to the nucleus of stem cells in the earliest stages of differentiation confirms once more the crucial role of NKX2-5 in steering the cardiomyocyte transcriptome.[140]

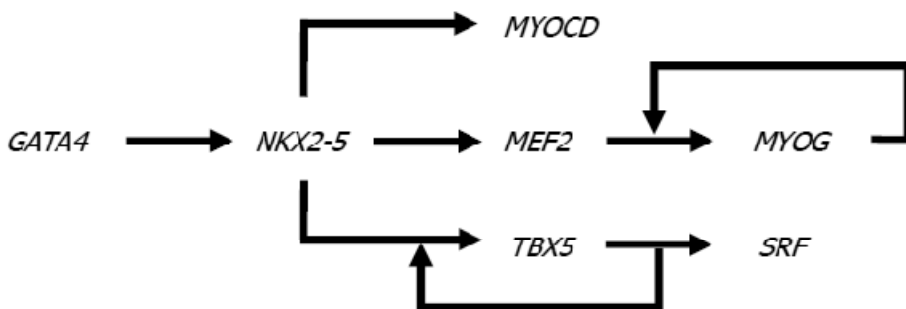


Figure 5. Regulatory roles amongst cardiac transcription factors.

Arrows indicate direct activating role on transcription through binding to the target gene promoter. The circulatory loops of myogenin and TBX5 reflect the self induction of these factors.

The more general SRF links together many elements of the cardiac transcriptional network and functions as the only known target for control of expression levels by the cardiac-enriched repressor HOP. Like MEF2, SRF connects the elements shared by the various muscle types with more cardiac-specific controllers and elements involved in controlling the core transcription machinery. Especially the ubiquitous

expressed p300 seems to be the principal factor for interacting with the core machinery.

There seems to be not a single general mechanism leading to cardiac-specific expression. The considerable variation in promoter structures of genes with a cardiac expression profile also becomes apparent from figure 1. Although it provides only visual information on possible binding sites for direct DNA binding proteins the figure clearly shows that the TFBS selected are highly enriched in the regulatory regions of cardiac genes compared to genes from other tissues, even including some of the closely related skeletal and smooth muscle. We described how many of the TFs involved in regulating cardiac-specific-transcription can act both through direct DNA binding and also as non-DNA binding co-factor, a relatively simple recording of the TFBs present within a promoter sequence thus only gives limited information on which TFs could possibly be involved in their regulation. The latter complicates an easy straightforward identification of all involved TFs through bioinformatics approaches demonstrating the need for experimental studies.

While the core machinery for the transcription activation of cardiac-specific expressed genes seems to be largely elucidated, knowledge of the promoters of many cardiac-specific genes is often non-existing or very limited at best. It is to be expected that further research in the regulation of these genes will result in the identification of transcriptional regulatory elements and factors that are yet unknown, and in functionally linking already known TFs to the cardiac-specific transcription machinery.

References

1. Weinmann, R., H.J. Raskas, and R.G. Roeder, Role of DNA-dependent RNA polymerases II and III in transcription of the adenovirus genome late in productive infection. *Proc Natl Acad Sci U S A*, 1974. 71(9): p. 3426-39.
2. Roeder, R.G., The role of general initiation factors in transcription by RNA polymerase II. *Trends in Biochemical Sciences*, 1996. 21(9): p. 327-335.
3. Muller, F., M.A. Demeny, and L. Tora, New Problems in RNA Polymerase II Transcription Initiation: Matching the Diversity of Core Promoters with a Variety of Promoter Recognition Factors. *J Biol Chem.*, 2007. 282(20): p. 14685-14689.
4. Juven-Gershon, T., et al., The RNA polymerase II core promoter -- the gateway to transcription. *Curr Opin Cell Biol.*, 2008. 20(3): p. 253-259.

5. Weis, L. and D. Reinberg, Transcription by RNA polymerase II: initiator-directed formation of transcription-competent complexes. *FASEB J.*, 1992. 6(14): p. 3300-3309.
6. Roeder, R.G., Transcriptional regulation and the role of diverse coactivators in animal cells. *FEBS Letters*, 2005. 579(4): p. 909-915.
7. Xu, M. and P.R. Cook, Similar active genes cluster in specialized transcription factories. *J Cell Biol.*, 2008. 181(4): p. 615-623.
8. Mitchell, J.A. and P. Fraser, Transcription factories are nuclear subcompartments that remain in the absence of transcription. *Genes Dev*, 2008. 22(1): p. 20-5.
9. Osborne, C.S., et al., Active genes dynamically colocalize to shared sites of ongoing transcription. *Nat Genet*, 2004. 36(10): p. 1065-71.
10. Schug, J., et al., Promoter features related to tissue specificity as measured by Shannon entropy. *Genome Biol*, 2005. 6(4): p. R33.
11. Roth, R.B., et al., Gene expression analyses reveal molecular relationships among 20 regions of the human CNS. *Neurogenetics*, 2006. 7(2): p. 67-80.
12. Sonnenberg, H., C.K. Chong, and A.T. Veress, Cardiac atrial factor--an endogenous diuretic? *Can J Physiol Pharmacol*, 1981. 59(12): p. 1278-9.
13. Nemeh, M.N. and J.P. Gilmore, Natriuretic activity of human and monkey atria. *Circ Res.*, 1983. 53(3): p. 420-423.
14. Gardner, D.G., et al., Extra-atrial expression of the gene for atrial natriuretic factor. *Proc Natl Acad Sci U S A*, 1986. 83(18): p. 6697-701.
15. Durocher, D., et al., The cardiac transcription factors Nkx2-5 and GATA-4 are mutual cofactors. *Embo J*, 1997. 16(18): p. 5687-96.
16. Bodmer, R., The gene tinman is required for specification of the heart and visceral muscles in *Drosophila*. *Development*, 1993. 118(3): p. 719-729.
17. Komuro, I. and S. Izumo, Csx: A Murine Homeobox-Containing Gene Specifically Expressed in the Developing Heart. *Proc Natl Acad Sci U S A*, 1993. 90(17): p. 8145-8149.
18. Lyons, I., et al., Myogenic and morphogenetic defects in the heart tubes of murine embryos lacking the homeo box gene Nkx2-5. *Genes Dev*, 1995. 9(13): p. 1654-66.
19. Lints, T., et al., Nkx-2.5: a novel murine homeobox gene expressed in early heart progenitor cells and their myogenic descendants. *Development*, 1993. 119(2): p. 419-431.
20. Lee, Y., et al., The Cardiac Tissue-Restricted Homeobox Protein Csx/Nkx2.5 Physically Associates with the Zinc Finger Protein GATA4 and Cooperatively Activates Atrial Natriuretic Factor Gene Expression. *Mol Cell Biol.*, 1998. 18(6): p. 3120-3129.
21. Durocher, D., et al., The atrial natriuretic factor promoter is a downstream target for Nkx- 2.5 in the myocardium. *Mol Cell Biol.*, 1996. 16(9): p. 4648-4655.
22. Chen, C.Y. and R.J. Schwartz, Identification of Novel DNA Binding Targets and Regulatory Domains of a Murine Tinman Homeodomain Factor, nkx-2.5. *J Biol Chem.*, 1995. 270(26): p. 15628-15633.

23. Sepulveda, J.L., et al., Combinatorial Expression of GATA4, Nkx2-5, and Serum Response Factor Directs Early Cardiac Gene Activity. *J Biol Chem.*, 2002. 277(28): p. 25775-25782.
24. Li, T., et al., Carboxyl Terminus of NKX2.5 Impairs its Interaction with p300. *J Mol Biol.*, 2007. 370(5): p. 976-992.
25. Arceci, R.J., et al., Mouse GATA-4: a retinoic acid-inducible GATA-binding transcription factor expressed in endodermally derived tissues and heart. *Mol Cell Biol*, 1993. 13(4): p. 2235-46.
26. Patient, R.K. and J.D. McGhee, The GATA family (vertebrates and invertebrates). *Curr Opin Genet Dev.*, 2002. 12(4): p. 416-422.
27. Zhao, R., et al., Loss of both GATA4 and GATA6 blocks cardiac myocyte differentiation and results in acardia in mice. *Dev Biol.*, 2008. 317(2): p. 614-619.
28. Charron, F., et al., Cooperative Interaction between GATA-4 and GATA-6 Regulates Myocardial Gene Expression. *Mol Cell Biol.*, 1999. 19(6): p. 4355-4365.
29. Morin, S., et al., GATA-dependent recruitment of MEF2 proteins to target promoters. *EMBO J.*, 2000. 19(9): p. 2046-2055.
30. Belaguli, N.S., et al., Cardiac tissue enriched factors serum response factor and GATA-4 are mutual coregulators. *Mol Cell Biol.*, 2000. 20(20): p. 7550-8.
31. Dai, Y.-S. and B.E. Markham, p300 Functions as a Coactivator of Transcription Factor GATA-4. *J Biol Chem.*, 2001. 276(40): p. 37178-37185.
32. Abraham, S.E., et al., p300, and p300-associated proteins, are components of TATA-binding protein (TBP) complexes. *Oncogene*, 1993. 8(6): p. 1639-47.
33. Sepulveda, J.L., et al., GATA-4 and Nkx-2.5 Coactivate Nkx-2 DNA Binding Targets: Role for Regulating Early Cardiac Gene Expression. *Mol Cell Biol.*, 1998. 18(6): p. 3405-3415.
34. Dai, Y.-S., et al., The Transcription Factors GATA4 and dHAND Physically Interact to Synergistically Activate Cardiac Gene Expression through a p300-dependent Mechanism. *J Biol Chem.*, 2002. 277(27): p. 24390-24398.
35. Fox, A.H., et al., Transcriptional cofactors of the FOG family interact with GATA proteins by means of multiple zinc fingers. *Embo J*, 1999. 18(10): p. 2812-22.
36. Dale, R.M., B.F. Remo, and E.C. Svensson, An alternative transcript of the FOG-2 gene encodes a FOG-2 isoform lacking the FOG repression motif. *Biochemical and Biophysical Research Communications*, 2007. 357(3): p. 683-687.
37. Rouf, R., et al., Increased FOG-2 in Failing Myocardium Disrupts Thyroid Hormone-Dependent SERCA2 Gene Transcription. *Circ Res.*, 2008. 103(5): p. 493-501.
38. Hu, X., et al., GATA4 regulates ANF expression synergistically with Sp1 in a cardiac hypertrophy model. *J Cell Mol Med*, 2010.

39. Lin, Q., et al., Control of Mouse Cardiac Morphogenesis and Myogenesis by Transcription Factor MEF2C. *Science*, 1997. 276(5317): p. 1404-1407.
40. Bassel-Duby, R., et al., Sequence elements required for transcriptional activity of the human myoglobin promoter in intact myocardium. *Circ Res.*, 1993. 73(2): p. 360-366.
41. Molkentin, J.D. and B.E. Markham, Myocyte-specific enhancer-binding factor (MEF-2) regulates alpha- cardiac myosin heavy chain gene expression *in vitro* and *in vivo*. *J Biol Chem.*, 1993. 268(26): p. 19512-19520.
42. Grayson, J., R. Bassel-Duby, and R.S. Williams, Collaborative interactions between MEF-2 and Sp1 in muscle-specific gene regulation. *J Cell Biochem.*, 1998. 70(3): p. 366-75.
43. Bassel-Duby, R., et al., Sequence elements required for transcriptional activity of the human myoglobin promoter in intact myocardium. *Circ Res.*, 1993. 73(2): p. 360-366.
44. Shore, P. and A.D. Sharrocks, The MADS-box family of transcription factors. *Eur J Biochem.*, 1995. 229(1): p. 1-13.
45. Feo, S., et al., Transcription of the human beta enolase gene (ENO-3) is regulated by an intronic muscle-specific enhancer that binds myocyte-specific enhancer factor 2 proteins and ubiquitous G-rich-box binding factors. *Mol Cell Biol.*, 1995. 15(11): p. 5991-6002.
46. Nurrish, S.J. and R. Treisman, DNA binding specificity determinants in MADS-box transcription factors. *Mol. Cell. Biol.*, 1995. 15(8): p. 4076-4085.
47. Messenguy, F. and E. Dubois, Role of MADS box proteins and their cofactors in combinatorial control of gene expression and cell development. *Gene*, 2003. 316: p. 1-21.
48. Cserjesi, P. and E.N. Olson, Myogenin induces the myocyte-specific enhancer binding factor MEF-2 independently of other muscle-specific gene products. *Mol Cell Biol.*, 1991. 11(10): p. 4854-4862.
49. Naya, F.J., et al., Transcriptional activity of MEF2 during mouse embryogenesis monitored with a MEF2-dependent transgene. *Development*, 1999. 126(10): p. 2045-2052.
50. Molkentin, J.D., et al., Cooperative activation of muscle gene expression by MEF2 and myogenic bHLH proteins. *Cell*, 1995. 83(7): p. 1125-1136.
51. Black, B.L. and E.N. Olson, Transcriptional control of muscle development by myocyte enhancer factor-2 (MEF2) proteins. *Annu Rev Cell Dev Biol.*, 1998. 14: p. 167-96.
52. Poizat, C., et al., Proteasome-mediated degradation of the coactivator p300 impairs cardiac transcription. *Mol Cell Biol.*, 2000. 20(23): p. 8643-54.
53. Sartorelli, V., et al., Molecular mechanisms of myogenic coactivation by p300: direct interaction with the activation domain of MyoD and with the MADS box of MEF2C. *Mol Cell Biol.*, 1997. 17(2): p. 1010-1026.
54. Zang, M.X., et al., Cooperative activation of atrial natriuretic peptide promoter by dHAND and MEF2C. *J Cell Biochem.*, 2004. 93(6): p. 1255-66.

55. Nojima, M., et al., The Positive Transcription Elongation Factor b Is an Essential Cofactor for the Activation of Transcription by Myocyte Enhancer Factor 2. *J Mol Biol.*, 2008. 382(2): p. 275-287.
56. Thomas, T., et al., The bHLH Factors, dHAND and eHAND, Specify Pulmonary and Systemic Cardiac Ventricles Independent of Left-Right Sidedness. *Dev Biol.*, 1998. 196(2): p. 228-236.
57. Cserjesi, P., et al., Expression of the Novel Basic Helix-Loop-Helix Gene eHAND in Neural Crest Derivatives and Extraembryonic Membranes during Mouse Development. *Dev Biol.*, 1995. 170(2): p. 664-678.
58. Natarajan, A., et al., Human eHAND, but not dHAND, is Down-regulated in Cardiomyopathies. *J Mol Cell Cardiol.*, 2001. 33(9): p. 1607-1614.
59. Russell, M.W., P. Baker, and S. Izumo, Cloning, chromosomal mapping, and expression of the human eHAND gene. *Mamm Genome*, 1997. 8(11): p. 863-5.
60. Hollenberg, S.M., et al., Identification of a new family of tissue-specific basic helix-loop-helix proteins with a two-hybrid system. *Mol Cell Biol.*, 1995. 15(7): p. 3813-3822.
61. Thattaliyath, B.D., B.A. Firulli, and A.B. Firulli, The basic-helix-loop-helix transcription factor HAND2 directly regulates transcription of the atrial natriuretic peptide gene. *J Mol Cell Cardiol.*, 2002. 34(10): p. 1335-44.
62. Ephrussi, A., et al., B lineage--specific interactions of an immunoglobulin enhancer with cellular factors *in vivo*. *Science*, 1985. 227(4683): p. 134-40.
63. Dai, Y.S. and P. Cserjesi, The basic helix-loop-helix factor, HAND2, functions as a transcriptional activator by binding to E-boxes as a heterodimer. *J Biol Chem*, 2002. 277(15): p. 12604-12.
64. Morin, S., et al., MEF2-dependent Recruitment of the HAND1 Transcription Factor Results in Synergistic Activation of Target Promoters. *J Biol Chem.*, 2005. 280(37): p. 32272-32278.
65. Smith, J., T-box genes: what they do and how they do it. *Trends in Genetics*, 1999. 15(4): p. 154-158.
66. Plageman, T.F., Jr. and K.E. Yutzey, Differential Expression and Function of Tbx5 and Tbx20 in Cardiac Development. 2004. p. 19026-19034.
67. Takeuchi, J.K., et al., Tbx5 specifies the left/right ventricles and ventricular septum position during cardiogenesis. *Development*, 2003. 130(24): p. 5953-5964.
68. Ghosh, T.K., et al., Physical Interaction between TBX5 and MEF2C Is Required for Early Heart Development. *Mol Cell Biol.*, 2009. 29(8): p. 2205-2218.
69. Georges, R., et al., Distinct Expression and Function of Alternatively Spliced Tbx5 Isoforms in Cell Growth and Differentiation. *Mol Cell Biol.*, 2008. 28(12): p. 4052-4067.
70. Bruneau, B.G., et al., A Murine Model of Holt-Oram Syndrome Defines Roles of the T-Box Transcription Factor Tbx5 in Cardiogenesis and Disease. *Cell*, 2001. 106(6): p. 709-721.

71. Hatcher, C.J., et al., Identification and localization of TBX5 transcription factor during human cardiac morphogenesis. *Dev Dyn*, 2000. 219(1): p. 90-5.
72. Minty, A. and L. Kedes, Upstream regions of the human cardiac actin gene that modulate its transcription in muscle cells: presence of an evolutionarily conserved repeated motif. *Mol Cell Biol*, 1986. 6(6): p. 2125-36.
73. Sartorelli, V., K. Webster, and L. Kedes, Muscle-specific expression of the cardiac alpha-actin gene requires MyoD1, CArG-box binding factor, and Sp1. *Genes Dev.*, 1990. 4(10): p. 1811-1822.
74. Biesiada, E., et al., Myogenic Basic Helix-Loop-Helix Proteins and Sp1 Interact as Components of a Multiprotein Transcriptional Complex Required for Activity of the Human Cardiac alpha -Actin Promoter. *Mol Cell Biol.*, 1999. 19(4): p. 2577-2584.
75. Weintraub, H., et al., The myoD gene family: nodal point during specification of the muscle cell lineage. *Science*, 1991. 251(4995): p. 761-6.
76. Weintraub, H., The MyoD family and myogenesis: Redundancy, networks, and thresholds. *Cell*, 1993. 75(7): p. 1241-1244.
77. Catala, F., et al., A skeletal muscle-specific enhancer regulated by factors binding to E and CArG boxes is present in the promoter of the mouse myosin light- chain 1A gene. *Mol Cell Biol.*, 1995. 15(8): p. 4585-4596.
78. Puri, P.L., et al., Differential Roles of p300 and PCAF Acetyltransferases in Muscle Differentiation. *Molecular Cell*, 1997. 1(1): p. 35-45.
79. Puri, P.L., et al., p300 is required for MyoD-dependent cell cycle arrest and muscle-specific gene transcription. *Embo J*, 1997. 16(2): p. 369-83.
80. Berkes, C.A., et al., Pbx Marks Genes for Activation by MyoD Indicating a Role for a Homeodomain Protein in Establishing Myogenic Potential. *Molecular Cell*, 2004. 14(4): p. 465-477.
81. Miner, J.H., J.B. Miller, and B.J. Wold, Skeletal muscle phenotypes initiated by ectopic MyoD in transgenic mouse heart. *Dev Biol.*, 1992. 114(4): p. 853-860.
82. Chen, C.Y., et al., Activation of the cardiac alpha-actin promoter depends upon serum response factor, Tinman homologue, Nkx-2.5, and intact serum response elements. *Dev Genet*, 1996. 19(2): p. 119-30.
83. Davis, R.L., et al., The MyoD DNA binding domain contains a recognition code for muscle-specific gene activation. *Cell*, 1990. 60(5): p. 733-746.
84. Miano, J.M., Cardiac-specific gene expression: a HANDful of factors. *J Mol Cell Cardiol*, 2002. 34(10): p. 1287-91.
85. Moss, J., T. McQuinn, and R. Schwartz, The avian cardiac alpha-actin promoter is regulated through a pair of complex elements composed of E boxes and serum response elements that bind both positive- and negative-acting factors. *J Biol Chem.*, 1994. 269(17): p. 12731-12740.
86. Groisman, R., et al., Physical Interaction between the Mitogen-responsive Serum Response Factor and Myogenic Basic-Helix-Loop-Helix Proteins. *J Biol Chem.*, 1996. 271(9): p. 5258-5264.

87. Weintraub, H., T. Genetta, and T. Kadesch, Tissue-specific gene activation by MyoD: determination of specificity by cis-acting repression elements. *Genes Dev.*, 1994. 8(18): p. 2203-2211.
88. Krainc, D., et al., Synergistic Activation of the N-Methyl-D-aspartate Receptor Subunit 1 Promoter by Myocyte Enhancer Factor 2C and Sp1. *J Biol Chem.*, 1998. 273(40): p. 26218-26224.
89. Chen, B. and R.W. Lim, Physical and Functional Interactions between the Transcriptional Inhibitors Id3 and ITF-2b. Evidence toward a novel Mechanism regulating Muscle-specific Gene Expression. *J Biol Chem.*, 1997. 272(4): p. 2459-2463.
90. Grueneberg, D.A., et al., Sequence-specific targeting of nuclear signal transduction pathways by homeodomain proteins. *Mol Cell Biol.*, 1995. 15(6): p. 3318-3326.
91. Chen, C.Y. and R.J. Schwartz, Recruitment of the tinman homolog Nkx-2.5 by serum response factor activates cardiac alpha-actin gene transcription. *Mol Cell Biol*, 1996. 16(11): p. 6372-84.
92. Mack, C.P., et al., Smooth muscle alpha-actin CArG elements coordinate formation of a smooth muscle cell-selective, serum response factor-containing activation complex. *Circ Res*, 2000. 86(2): p. 221-32.
93. Molinari, S., et al., A Novel Complex Regulates cardiac actin Gene Expression through Interaction of Emb, a Class VI POU Domain Protein, MEF2D, and the Histone Transacetylase p300. *Mol Cell Biol.*, 2004. 24(7): p. 2944-2957.
94. Lemonnier, M. and M.E. Buckingham, Characterization of a cardiac-specific enhancer, which directs alpha-cardiac actin gene transcription in the mouse adult heart. *J Biol Chem.*, 2004. 279(53): p. 55651-8.
95. Creemers, E.E., et al., Coactivation of MEF2 by the SAP Domain Proteins Myocardin and MASTR. *Molecular Cell*, 2006. 23(1): p. 83-96.
96. Wang, D., et al., Activation of cardiac gene expression by myocardin, a transcriptional cofactor for serum response factor. *Cell*, 2001. 105(7): p. 851-62.
97. Chen, J., et al., Myocardin: a component of a molecular switch for smooth muscle differentiation. *J Mol Cell Cardiol*, 2002. 34(10): p. 1345-56.
98. Yoshida, T., et al., Myocardin Is a Key Regulator of CArG-Dependent Transcription of Multiple Smooth Muscle Marker Genes. *Circ Res*, 2003. 92(8): p. 856-864.
99. Long, X., et al., Myocardin is sufficient for a smooth muscle-like contractile phenotype. *Arterioscler Thromb Vasc Biol*, 2008. 28(8): p. 1505-10.
100. Ueyama, T., et al., Myocardin Expression Is Regulated by Nkx2.5, and Its Function Is Required for Cardiomyogenesis. *Mol Cell Biol.*, 2003. 23(24): p. 9222-9232.
101. van Tuyn, J., et al., Activation of cardiac and smooth muscle-specific genes in primary human cells after forced expression of human myocardin. *Cardiovasc. Res.*, 2005. 67(2): p. 245-255.

102. Hinson, J.S., et al., Smooth muscle cell-specific transcription is regulated by nuclear localization of the myocardin-related transcription factors. *Am J Physiol Heart Circ Physiol*, 2007. 292(2): p. H1170-1180.
103. Chang, P.S., et al., Muscle specificity encoded by specific serum response factor-binding sites. *J Biol Chem.*, 2001. 276(20): p. 17206-12.
104. Izumo, S. and V. Mahdavi, Thyroid hormone receptor alpha-isoforms generated by alternative splicing differentially activate myosin HC gene transcription. *Nature*, 1988. 334(6182): p. 539-542.
105. Sayen, M.R., D.K. Rohrer, and W.H. Dillmann, Thyroid hormone response of slow and fast sarcoplasmic reticulum Ca²⁺ ATPase mRNA in striated muscle. *Mol Cell Endocrinol.*, 1992. 87(1-3): p. 87-93.
106. Lee, Y., et al., Myocyte-specific enhancer factor 2 and thyroid hormone receptor associate and synergistically activate the alpha-cardiac myosin heavy- chain gene. *Mol Cell Biol.*, 1997. 17(5): p. 2745-2755.
107. Lee, T.C., et al., Activation of skeletal alpha-actin gene transcription: the cooperative formation of serum response factor-binding complexes over positive cis-acting promoter serum response elements displaces a negative-acting nuclear factor enriched in replicating myoblasts and nonmyogenic cells. *Mol Cell Biol*, 1991. 11(10): p. 5090-100.
108. Yao, Y.-L., W.-M. Yang, and E. Seto, Regulation of Transcription Factor YY1 by Acetylation and Deacetylation. *Mol Cell Biol.*, 2001. 21(17): p. 5979-5991.
109. Chen, C.-Y. and R.J. Schwartz, Competition Between Negative Acting YY1 versus Positive Acting Serum Response Factor and Tinman Homologue Nkx-2.5 Regulates Cardiac alpha-Actin Promoter Activity. *Mol Endocrinol.*, 1997. 11(6): p. 812-822.
110. Sucharov, C.C., et al., Yin Yang 1 Is Increased in Human Heart Failure and Represses the Activity of the Human alpha-Myosin Heavy Chain Promoter. *J Biol Chem.*, 2003. 278(33): p. 31233-31239.
111. Hamamori, Y., et al., The basic domain of myogenic basic helix-loop-helix (bHLH) proteins is the novel target for direct inhibition by another bHLH protein, Twist. *Mol Cell Biol.*, 1997. 17(11): p. 6563-6573.
112. Hamamori, Y., et al., Regulation of Histone Acetyltransferases p300 and PCAF by the bHLH Protein Twist and Adenoviral Oncoprotein E1A. *Cell*, 1999. 96(3): p. 405-413.
113. Shirvani, S., et al., CHF1/Hey2 suppresses SM-MHC promoter activity through an interaction with GATA-6. *Biochem Biophys Res Commun*, 2006. 339(1): p. 151-6.
114. Fischer, A., et al., Hey basic helix-loop-helix transcription factors are repressors of GATA4 and GATA6 and restrict expression of the GATA target gene ANF in fetal hearts. *Mol Cell Biol*, 2005. 25(20): p. 8960-70.
115. Leimeister, C., et al., Hey genes: a novel subfamily of hairy- and Enhancer of split related genes specifically expressed during mouse embryogenesis. *Mech Dev*, 1999. 85(1-2): p. 173-7.

116. Chin, M.T., et al., Cardiovascular Basic Helix Loop Helix Factor 1, a Novel Transcriptional Repressor Expressed Preferentially in the Developing and Adult Cardiovascular System. *J Biol Chem.*, 2000. 275(9): p. 6381-6387.
117. Koibuchi, N. and M.T. Chin, CHF1/Hey2 Plays a Pivotal Role in Left Ventricular Maturation Through Suppression of Ectopic Atrial Gene Expression. *Circ Res.*, 2007. 100(6): p. 850-855.
118. Yamagishi, H., et al., The Combinatorial Activities of Nkx2.5 and dHAND Are Essential for Cardiac Ventricle Formation. *Developmental Biology*, 2001. 239(2): p. 190-203.
119. Bao, Z.Z., et al., Regulation of chamber-specific gene expression in the developing heart by Irx4. *Science*, 1999. 283(5405): p. 1161-4.
120. Xu, J., et al., Myocyte Enhancer Factors 2A and 2C Induce Dilated Cardiomyopathy in Transgenic Mice. *J Biol Chem.*, 2006. 281(14): p. 9152-9162.
121. McKinsey, T.A., C.L. Zhang, and E.N. Olson, Control of muscle development by dueling HATs and HDACs. *Curr Opin Genet Dev.*, 2001. 11(5): p. 497-504.
122. McKinsey, T.A. and E.N. Olson, Regulation of Muscle Gene Expression by Histone Deacetylases, in *Histone deacetylases, Transcriptional Regulation and Other Cellular Functions*, E. Verdin, Editor. 2006, Humana Press Inc.: Totowa, NJ U.S.A. p. 107-128.
123. Shin, C.H., et al., Modulation of Cardiac Growth and Development by HOP, an Unusual Homeodomain Protein. *Cell*, 2002. 110(6): p. 725-735.
124. Chen, F., et al., Hop is an unusual homeobox gene that modulates cardiac development. *Cell*, 2002. 110(6): p. 713-23.
125. Kim, T.-g., et al., Jumonji Represses Atrial Natriuretic Factor Gene Expression by Inhibiting Transcriptional Activities of Cardiac Transcription Factors. *Mol Cell Biol.*, 2004. 24(23): p. 10151-10160.
126. Bovill, E., et al., Induction by left ventricular overload and left ventricular failure of the human Jumonji gene (JARID2) encoding a protein that regulates transcription and reexpression of a protective fetal program. *J Thorac Cardiovasc Surg.*, 2008. 136(3): p. 709-716.
127. Kim, T.-g., et al., Jumonji represses alpha-cardiac myosin heavy chain expression via inhibiting MEF2 activity. *Biochem Biophys Res Commun.*, 2005. 329(2): p. 544-553.
128. Lu, J.-r., et al., FOG-2, a Heart- and Brain-Enriched Cofactor for GATA Transcription Factors. *Mol Cell Biol.*, 1999. 19(6): p. 4495-4502.
129. Svensson, E.C., et al., Molecular cloning of FOG-2: A modulator of transcription factor GATA-4 in cardiomyocytes. *Proc Natl Acad Sci U S A*, 1999. 96(3): p. 956-961.
130. Svensson, E.C., et al., A Functionally Conserved N-terminal Domain of the Friend of GATA-2 (FOG-2) Protein Represses GATA4-Dependent Transcription. *J Biol Chem.*, 2000. 275(27): p. 20762-20769.
131. Lien, C., et al., Control of early cardiac-specific transcription of Nkx2-5 by a GATA-dependent enhancer. *Development*, 1999. 126(1): p. 75-84.

132. Brown, C.O., III, et al., The Cardiac Determination Factor, Nkx2-5, Is Activated by Mutual Cofactors GATA-4 and Smad1/4 via a Novel Upstream Enhancer. *J Biol Chem.*, 2004. 279(11): p. 10659-10669.
133. Dodou, E., et al., Mef2c is a direct transcriptional target of ISL1 and GATA factors in the anterior heart field during mouse embryonic development. *Development*, 2004. 131(16): p. 3931-3942.
134. Edmondson, D.G., et al., Analysis of the myogenin promoter reveals an indirect pathway for positive autoregulation mediated by the muscle-specific enhancer factor MEF-2. *Mol Cell Biol.*, 1992. 12(9): p. 3665-3677.
135. Molkentin, J.D. and E.N. Olson, Combinatorial control of muscle development by basic helix-loop-helix and MADS-box transcription factors. *Proc Natl Acad Sci U S A*, 1996. 93(18): p. 9366-9373.
136. Barron, M.R., et al., Serum response factor, an enriched cardiac mesoderm obligatory factor, is a downstream gene target for TBX genes. *J Biol Chem.*, 2004: p. M412408200.
137. Maves, L., et al., Pbx acts with Hand2 in early myocardial differentiation. *Dev Biol.*, 2009. 333(2): p. 409-18.
138. Takeuchi, J.K. and B.G. Bruneau, Directed transdifferentiation of mouse mesoderm to heart tissue by defined factors. *Nature*, 2009. 459(7247): p. 708-711.
139. Ieda, M., et al., Direct Reprogramming of Fibroblasts into Functional Cardiomyocytes by Defined Factors. *Cell*, 2010. 142(3): p. 375-386.
140. Arminan, A., et al., Cardiac transcription factors driven lineage-specification of adult stem cells. *J Cardiovasc Transl Res*, 2010. 3(1): p. 61-5.

CHAPTER 3

CREATION AND VALIDATION OF A SET OF CARDIOVASCULAR PROMOTER REPORTER CONSTRUCTS

Ruben S.R.M. Martherus; Erika D.J. Timmer; Sabina J.V. Vanherle;
Torik A.Y. Ayoubi

Submitted

Abstract

With the human genome decoded, one of the major tasks remaining is that of identifying the role of the non-coding elements. A considerable part of these non-coding elements are promoter sequences, involved in regulating transcription, the process in which the DNA is transcribed to the intermediate mRNA. For many genes little to nothing is known about their promoter sequences. Knowledge of their promoter architecture would however be desirable in order to understand the genes' regulation in both health and disease. In this study we investigated the human promoters of a number of genes with a more cardiac-enriched expression profile, cysteine and glycine rich protein 3 (*CSRP3*), telethonin/ titin capping protein (*TCAP*), myosin light chain 3 (*MYL3*) and myosin light chain 7 (*MYL7*). Based on cross-species conservation and the occurrence of elements with known (heart-) muscle regulatory properties we selected and cloned their immediate 5'-flanking sequences of the genes. Experimental validation using luciferase reporter assays show that the tissue specificity of the selected fragments correlates with that of the expression profile from the endogenous genes.

Introduction

While for many genes the DNA sequences encoding the eventual proteins are known and often even functionally described into great detail, the characterization of the promoters is often lagging behind. Since transcription of a gene is the prelude to the eventual formation of a functional protein, research into the orchestration of a gene's transcription can greatly assist in understanding how a gene functions. All this also holds true for many genes with a cardiac-enriched expression profile. The protein encoding sequences of many cardiac-specific genes have been studied in detail, often in relation to mutations leading to hereditary heart-diseases.[1, 2] Meanwhile, somewhat surprisingly, the promoters of many of these genes have not been investigated at all. More knowledge on the role of regulatory elements within the promoter sequences and the transcription factors (TFs) associating with them could provide a better understanding of the

regulation of gene activity in the heart in both health and disease. Knowing which genes are targeted by a specific TF can for example aid in understanding cardiac disorders such as the Holt-Oram syndrome, which pathogenesis originates from mutations in the key cardiac transcription factor TBX5.[2]

Identification of relevant promoter elements will eventually require some form of experimental validation. Luciferase reporter vectors have become one of the major tools for experimental studies on promoters. In the study presented here we described the partial identification of the promoter sequences of a number of genes involved in cardiomyocytes and their validation using transient transfection of luciferase reporter constructs in various types of muscle cells. Determining the tissue specificity of promoter fragments is important. Closely related types of tissues such as cardiac, skeletal and smooth muscle share many similarities, including some of the mechanisms involved in transcription. A good example is cardiac α -actin (*ACTC1*). For this gene, the proximal promoter region -110 to +68 is sufficient for restricting transcription to the striated muscles.[3] For the eventual cardiac-specific transcription of *ACTC1*, additional promoter-sequences are required, including an enhancer located within -2.353 to -1.36 kbp relative to the TSS.[4] For many genes with a cardiomyocyte specific expression profile it seems that a reasonably short stretch of the proximal promoter is often sufficient for driving their cardiomyocyte-specific transcription. For natriuretic peptide precursor type A (*NPPA/ANF*) and muscle α -myosin heavy chain (*MYH6*) respectively 700 bp and 344 bp of the 5'proximal sequences are sufficient for cardiac-specific expression.[5, 6] We therefore limited our search to the immediate 5' flanking sequences.

We present here the partial elucidation of promoter regions for the genes: cysteine and glycine-rich protein 3 (*CSRP3*), lysyl oxidase-like 1 (*LOXL1*), myosin light chain 3 (*MYL3*), myosin light chain 7 (*MYL7*), heat shock 27 kDa protein family member-7 (*HSPB7*) and titin-capping protein / telethonin (*TCAP*). For some of the genes included in this study, little to none information had been available on their promoter so far, of others; previously published data was only available for the mouse promoters. In the study we present here we provide data on the transcriptional activity of the human promoters and demonstrated the tissue specificity of selected promoter fragments in various muscle tissues including cardiomyocytes.

Materials and methods

Creation of promoter constructs

Promoter sequences of interest were cloned into pGL4.10 (Promega, Madison, WI USA) using standard molecular biology techniques. The promoter regions were obtained via PCR on human DNA. Clones were verified through DNA sequencing using the BigDye cycle sequencing system (Applied Biosystems, Carlsbad CA USA).

Cell culture and transfection

Primary culture of neonatal rat ventricular myocytes: Primary cultures of neonatal rat ventricular cardiomyocytes were generated and prepared as described earlier.[7] C2C12 (mouse skeletal myoblasts) and A7R5 (rat aortic smooth muscle cells) cultures were obtained from ATCC (Manassas, VA USA). C2C12 cells were maintained in Dulbecco's modified Eagle Medium, DMEM, 10% fetal bovine serum (FBS) and transferred to DMEM supplemented with 2% horse serum to induce and maintain differentiation. A7R5 cells were maintained in DMEM supplemented with 15% FBS, and for differentiation cells were grown in DMEM supplemented with 2% horse serum.

Transfections

Cells were transfected with plasmid using FugeneHD (Roche) according to manufacturer's instructions. Transfection mixtures were composed to equal molar amounts of all promoter constructs. To adjust total DNA per transfection without introducing additional luciferase encoding DNA, 'empty' pGEM-Teasy (Promega, Madison, WI USA) was used. Transfection experiments were performed in 48 well plates, using 1 μ l Fugene HD and 0.33 μ g total DNA per well. To enable correction for transfection efficiency, cells were co-transfected with 10 ng of pGL4.75, Renilla Luciferase under control of the CMV promoter (Promega, Madison, WI USA), per well.

Results

Selection of promoter fragments

Based on the public available databases of SymAtlas (later replaced by BioGPS), the Reference database for gene expression analysis, RefExA (http://157.82.78.238/refexa/main_search.jsp), NCBI Gene Expression Omnibus (GEO) and the Database of Transcription start sites (DBTSS) we identified the relevant transcripts.[8-10] Genomic DNA sequences were retrieved via the Biomart website. Information deposited in the DBTSS combined with data from the NCBI Refseq database was used to determine the transcription start site of the chosen transcripts.[8, 9] This precise annotation of the transcription start site (TSS) is of great importance since genes can encode for different transcripts via the use of alternative promoters, which can be located at great distances from one another.[11] In relation to this DBTSS was very useful since it provides information on the TSS of full length cDNA clones obtained from samples from different tissues, including some representing the transcripts in various types of muscle. Also, attention was given to the presence of core transcription regulating elements such as TATA-boxes and INR-elements.

Because of technical limitations such as transfection efficiency we limited the regions to be cloned to about 5 kb. Therefore, when screening for CNSs we focused on the region up to 5kb from the transcription start site. The screening for CNSs was done by comparing the human sequences with those of mice, rat, dog and cow in the genome-VISTA program.[12] The results of these alignments are provided as on-line available supplementary data.

For the *in vitro* studies we focused on the strongly conserved regions (>70% homology with mouse, rat, dog or cow) directly upstream of the TSS. However in the case of *MYL7* we observed also a very strong conservation in the first intron. To investigate the occurrence of crucial regulatory elements within this region we generated two clones, one including the transcribed region up to and including the second exon. Notion was also given to any non-conserved clusters of TFBs with a known involvement in regulating transcription in muscle.[13] For this we screened for matches to consensus binding site sequences as described by Xie et al.[14] and sequences available in the TRANSFAC database.[15] In the case of *HSPB7*, various candidate TSS were found, to include them all, we selected our cut off for this promoter fragment further downstream of the TSS presented by the Refseq NM_014424

mRNA. The selected promoter regions of the genes investigated are given in table 1.

Validation of reporter constructs

Of all the genes investigated, we cloned the human DNA-derived promoter regions in the pGL4.10 luciferase reporter vector. Of the genes *CSRP3* and *MYL7* two regions of varying length were selected. The proposed *CSRP3* promoter region contains a short CNSs stretching to -780 bp. However, several additional potentially crucial binding sites were identified further upstream, but still within such proximity that the combined region could still be cloned as one contiguous fragment. The first intron of *MYL7* was found to contain a considerable amount of CNSs. To investigate their potential activity, an additional *MYL7* construct was generated containing the genomic region up to and including the second exon. Care was given to keep coding regions in frame of the ATG start codon located in the first exon of *MYL7*.

The transcriptional activity and tissue specificity of the generated constructs were analyzed through transfections into cardiomyocytes, cell lines representative for skeletal muscle, smooth muscle and HELA cells (figure 1). For comparison of the tissue specificity of the genes we also performed transfections with a construct containing the smooth muscle-specific lysyl oxidase (*LOXL1*) promoter and with one bearing the promoter of alpha myosin heavy chain (*MYH6*).[16, 17]

The promoter fragments of *MYL3* and *MYL7* display a high degree of tissue specificity for cardiomyocytes. While the shorter *CSRP3* promoter fragment displays high transcriptional activity in all the different muscle cells, the larger *CSRP3* fragment yields a more tissue restricted expression profile. The activity measured for the selected *HSPB7* promoter fragment did not rise significantly above that of the empty pGL4.10 control vector in any of the cell-lines tested.

Table 1. Promoter fragments

HGNC symbol	Promoter Region ^a bp	DBTSS ^b	Refseq ^c	Tata-box ^d	Previously described
CSRP3	-4923 / +77 -780 / +77	+		+	-
FABP3	-1247 / +40		NM_004102	+	mouse[13]
HSPB7	-3888 / +392	+			-
LOXL1	-4080 / +14	+			rat[14], human[15]
MYBPC1	-3068 / +72	+			-
MYH6	-2028 / +30		NM_002471		human[16], rat[6]
MYL3	-1951 / +64	+	NM_000258		-
MYL7	-1038 / +71 -1038 / +329	+	NM_021223	+	zebrafish[17]
TCAP	-637 / +23	+	NM_003673		-

^aPromoter regions incorporated in luciferase reporter vectors used in this study. ^b'+' indicates where the selection of the TSS was supported by data from the DBTSS database. ^cWhen the selected TSS coincided with that of a known Refseq transcript, the corresponding Refseq ID is indicated. ^dBinding site matching TATA-box consensus sequence was present at -25 to -30 bp relative to selected TSS.

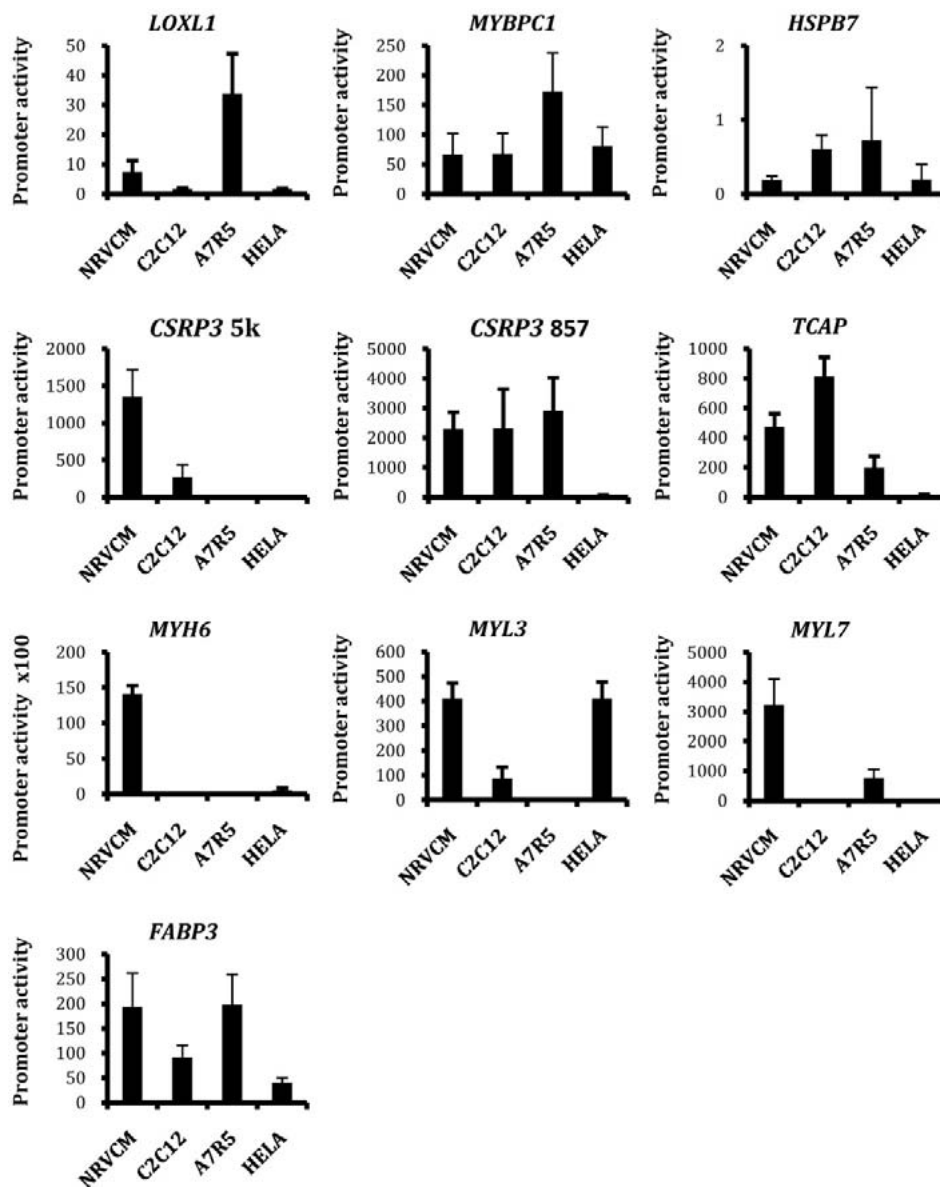


Figure 1. Transcriptional activity of promoter fragments.

Promoter reporter constructs derived from human promoter sequences were transfected into primary cultures of neonatal rat ventricular cardiomyocytes (NRVCN), mouse skeletal muscle myocytes (C2C12), rat aortic smooth muscle cells (A7R5) and human cervical cancer cells (HELA). Promoter activity was calculated on fold activation over the background obtained with equimolar amounts of the empty pGL4.10 reporter vector.

Discussion

For many genes with a cardiac-specific expression little to none information is available on their promoter architecture. To assist in the advance on research on some of these genes we investigated the 5'-flanking sequences and generated and validated a set of luciferase reporter vector constructs.

We created a *LOXL1* reporter construct that, unlike Csiszar et al, did not include the first intron.[16] Furthermore, the region we incorporated was 4 kb where Csiszar et al. included only up to -925 bp of the 5'-flanking region. The observed expression profile of the resulting *LOXL1* promoter fragment corresponded to what is observed for the endogenous gene in for example the RefExA database. Interestingly, hardly any expression was observed in HELA cells while in the study by Csiszar et al. expression was observed in the osteosarcoma derived cell-line. The different origins of both cancer-cell lines as well as the difference in the selected promoter could be causative for this.

Based on the gene expression databases we expected a strong enrichment of *MYBPC1* expression in the skeletal muscle cells. The strongest activity of our luciferase reporter construct was however observed in A7R5 cells for which *MYBPC1* expression was not expected. A recent study showed that *MYBPC1* expression in cultured skeletal muscle cells was strongly down-regulated when compared to normal muscle tissue.[18] This explains for the low luciferase-levels in C2C12 cells but not for the occurrence of expression in smooth muscle cells and cardiomyocytes. Possibly some sequences providing inhibitory activity within these tissues are located outside of the promoter region selected here.

Like with many other cardiac-specific genes, relatively short stretches of the immediate 5'-flanking sequences appear to be sufficient for driving transcription in cardiomyocytes of *CSRP3*, *MYL3*, *MYL7* and *TCAP*. The investigated fragment of *MYL3* displays high levels of activity in cardiomyocytes and some activity in C2C12 cells. This corresponds to gene expression data from the endogenous human gene, which is known to be expressed in both heart and skeletal muscle (see for example the REFEXA database).

For *MYL7*, so far, only the zebrafish promoter has been described before.[19] We began here to provide some functional data on the human promoter. Although *MYL7* is principally expected to be a fully cardiac-specific gene, the selected *MYL7* promoter fragment still displayed some residual activity in A7R5 cells while it did not display

any activity in C2C12 cells. This is somewhat remarkable since skeletal muscle cells, like cardiomyocytes, are striated muscle cells, and thus structurally and functionally more related.

Additional studies would be required to verify whether these observations are resulting from differences in the endogenous transcriptional program between A7R5 cell cultures and the smooth muscle *in vivo*, or from regulatory elements located outside the promoter regions incorporated in this study.

TCAP expression is normally observed in all three types of muscle.[20, 21] The selected promoter fragment indeed displayed activity in all three types of muscle cells, while at the same time no activity was observed in HELA cells. It therefore seems that a relatively short promoter fragment constitutes the expression profile observed *in vivo*.

For FABP3, so far only the mouse promoter had been characterized *in vitro*. [22] Similar to what was found for the mouse promoter we found the investigated human FABP3 promoter-fragment to be active in both cardiomyocytes and C2C12 cells as well as smooth muscle. The promoter fragment thus yields an expression profile corresponding to what has been found for the FABP3 protein.[23]

We found for *CSRP3* that the larger promoter-fragment displayed much lower expression levels but at the same time led to a more tissue restricted expression profile. Although the gene has been found to be expressed in blood vessels we failed to detect promoter activity of the larger promoter fragment in cultured smooth muscle of aortic origin.[24] A possible explanation could be the influence of stretch. Campos and co-workers found that expression of *CSRP3* was high in smooth muscle cells subjected to stretch, while no expression was detected when the cells were not subjected to stretching. Such a stretch dependent activation could also explain why the *CSRP3* promoter activity was higher in the cardiomyocytes. Under the culture conditions employed in our study, the cardiomyocytes display frequent spontaneous contractions. It therefore could also well be that the 5kb *CSRP3*-promoter fragment becomes activated in A7R5 and C2C12 cells if these cells were to be subjected to stretch. If so, further investigation of this *CSRP3* promoter fragment could provide valuable insights into the mechanisms underlying this response.

Conclusion

We provided here for the first time information on the tissue-specific transcriptional activity of a number of genes expressed in cardiomyocytes, skeletal and smooth muscle. In general, the immediate conserved non-coding 5'-flanking sequences of genes with a cardiac-enriched expression profile appear to represent the vast majority of the regulatory elements involved in the regulation of these genes. To confirm specificity for a specific type of muscle, *in vitro* models such as cultures of C2C12 and A7R5 cells can potentially provide valuable information. However, some limitations in using cell cultures for studying transcriptional mechanisms became apparent. The variations in expression of for example MYBPC1 and CSRP3 in cultured cells compared to the *in vivo* situation provided good examples for this. More definitive confirmation of the promoter fragments investigated here could come from *in vivo* studies. Alternatively, muscle cell cultures might be subjected to additional treatments such as exposure to stretch in order to make them a more representative and reliable model for transcriptional research *in vitro*.

Acknowledgements

This study was financed by a grant from the Transnationale Universiteit Limburg

References

1. P. Richard, P. Charron, L. Carrier, et al., Hypertrophic Cardiomyopathy: Distribution of Disease Genes, Spectrum of Mutations, and Implications for a Molecular Diagnosis Strategy. *Circulation*, 107 (2003) 2227-2232.
2. B.G. Bruneau, G. Nemer, J.P. Schmitt, et al., A Murine Model of Holt-Oram Syndrome Defines Roles of the T-Box Transcription Factor Tbx5 in Cardiogenesis and Disease. *Cell*, 106 (2001) 709-721.
3. A. Minty and L. Kedes, Upstream regions of the human cardiac actin gene that modulate its transcription in muscle cells: presence of an evolutionarily conserved repeated motif. *Mol Cell Biol*, 6 (1986) 2125-36.
4. M. Lemonnier and M.E. Buckingham, Characterization of a cardiac-specific enhancer, which directs alpha-cardiac actin gene transcription in the mouse adult heart. *J Biol Chem.*, 279 (2004) 55651-8.
5. D. Durocher, F. Charron, R. Warren, et al., The cardiac transcription factors Nkx2-5 and GATA-4 are mutual cofactors. *Embo J*, 16 (1997) 5687-96.

6. J.D. Molkentin, S.M. Jobe, and B.E. Markham, alpha-myosin Heavy Chain Gene Regulation: Delineation and Characterization of the Cardiac Muscle-specific Enhancer and Muscle-specific Promoter. *J Mol Cell Cardiol.*, 28 (1996) 1211-1225.
7. R.S.R.M. Martherus, V.A. Zeijlemaker, and T.A.Y. Ayoubi, Electrical stimulation of primary neonatal rat ventricular cardiomyocytes using pacemakers. *Biotechniques* 48 (2010) 65-67
8. A.I. Su, T. Wiltshire, S. Batalov, et al., A gene atlas of the mouse and human protein-encoding transcriptomes. *PNAS*, 101 (2004) 6062-6067.
9. T. Barrett, T.O. Suzek, D.B. Troup, et al., NCBI GEO: mining millions of expression profiles--database and tools. *Nucl Acids Res.*, 33 (2005) D562-566.
10. Y. Suzuki, R. Yamashita, K. Nakai, et al., DBTSS: DataBase of human Transcriptional Start Sites and full-length cDNAs, in *Nucl Acids Res.* 2002. p. 328-31.
11. T. Ayoubi and W. Van De Ven, Regulation of gene expression by alternative promoters. *FASEB J.*, 10 (1996) 453-460.
12. O. Couronne, A. Poliakov, N. Bray, et al., Strategies and Tools for Whole-Genome Alignments. 2003. p. 73-80.
13. R.M. Cripps and E.N. Olson, Control of Cardiac Development by an Evolutionarily Conserved Transcriptional Network. *Dev Biol.*, 246 (2002) 14-28.
14. X. Xie, J. Lu, E.J. Kulbokas, et al., Systematic discovery of regulatory motifs in human promoters and 3[prime] UTRs by comparison of several mammals. *Nature*, 434 (2005) 338-345.
15. E. Wingender, X. Chen, E. Fricke, et al., The TRANSFAC system on gene expression regulation. *Nucl. Acids Res.*, 29 (2001) 281-283.
16. K. Csiszar, I. Entersz, P.C. Trackman, et al., Functional analysis of the promoter and first intron of the human lysyl oxidase gene. *Mol Biol Rep*, 23 (1996) 97-108.
17. S. Gao, Y. Zhao, L. Kong, et al., Cloning and Characterization of the Rat Lysyl Oxidase Gene Promoter. *J Biol Chem.*, 282 (2007) 25322-25337.
18. F. Raymond, S. Metairon, M. Kussmann, et al., Comparative gene expression profiling between human cultured myotubes and skeletal muscle tissue. *BMC Genomics*, 11 (2010) 125.
19. L. Maves, A. Tyler, C.B. Moens, et al., Pbx acts with Hand2 in early myocardial differentiation. *Dev Biol.*, 333 (2009) 409-18.
20. A. Mazzone, P.R. Strege, D.J. Tester, et al., A Mutation in Telethonin Alters Nav1.5 Function. *J Biol Chem.*, 283 (2008) 16537-16544.
21. G. Valle, G. Faulkner, A. De Antoni, et al., Telethonin, a novel sarcomeric protein of heart and skeletal muscle. *FEBS Letters*, 415 (1997) 163-168.
22. Q. Qian, L. Kuo, Y.-T. Yu, et al., A Concise Promoter Region of the Heart Fatty Acid-Binding Protein Gene Dictates Tissue-Appropriate Expression. *Circ Res*, 84 (1999) 276-289.

23. W. Zschesche, A.H. Kleine, E. Spitzer, et al., Histochemical localization of heart-type fatty-acid binding protein in human and murine tissues. *Histochem Cell Biol*, 103 (1995) 147-56.
24. L.C.G. Campos, A.A. Miyakawa, V.G. Barauna, et al., Induction of CRP3/MLP expression during vein arterialization is dependent on stretch rather than shear stress. *Cardiovasc Res*, 83 (2009) 140-147.

Supplementary Material

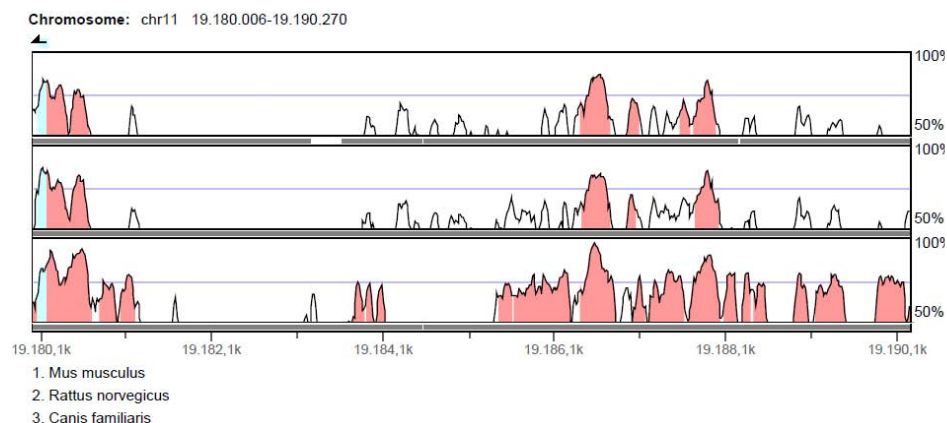
Supplementary figures.

Orthologues sequences alignments.

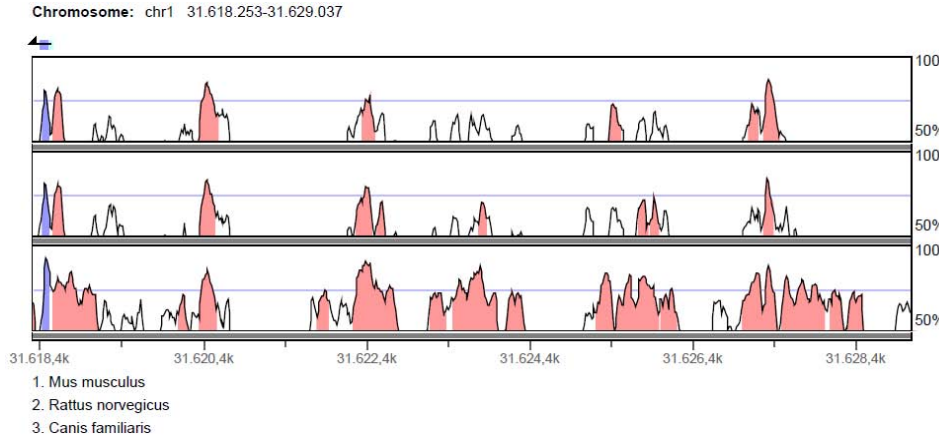
To identify conserved non-coding sequences, using the Vista genome browser (<http://pipeline.lbl.gov/vista/index.shtml>) alignments were generated of the human genomic sequences with those of mouse, rat and dog for up to 10kb 5' upstream of the transcription start site.

Annotations:  Gene  Untranslated region
 Exon  Conserved con-coding sequence

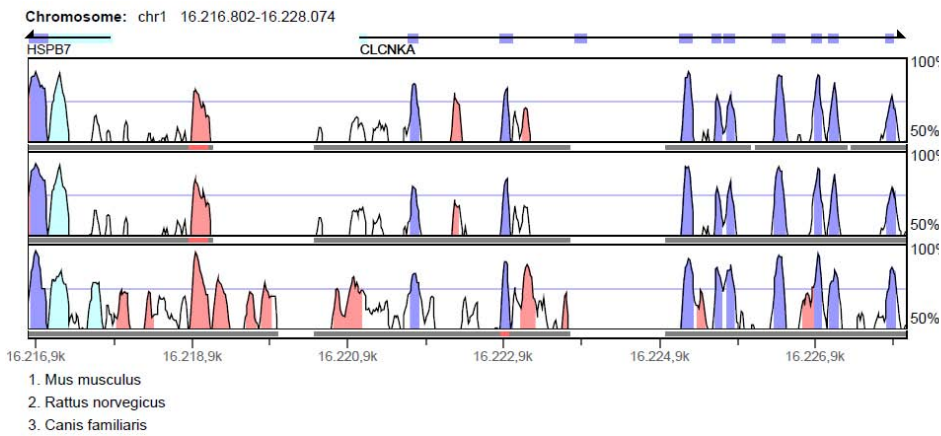
CSRP3



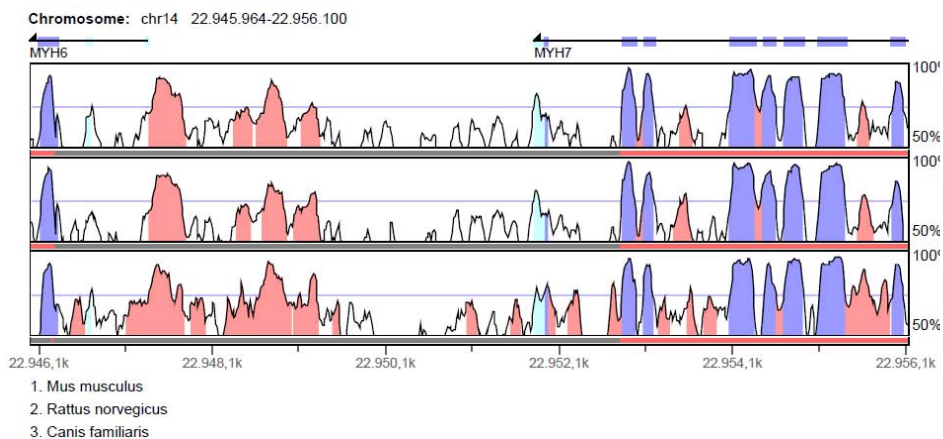
FABP3



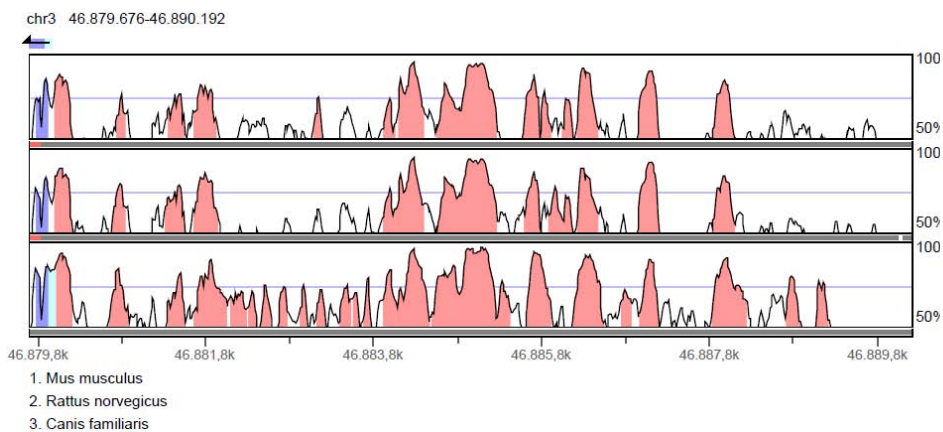
HSPB7



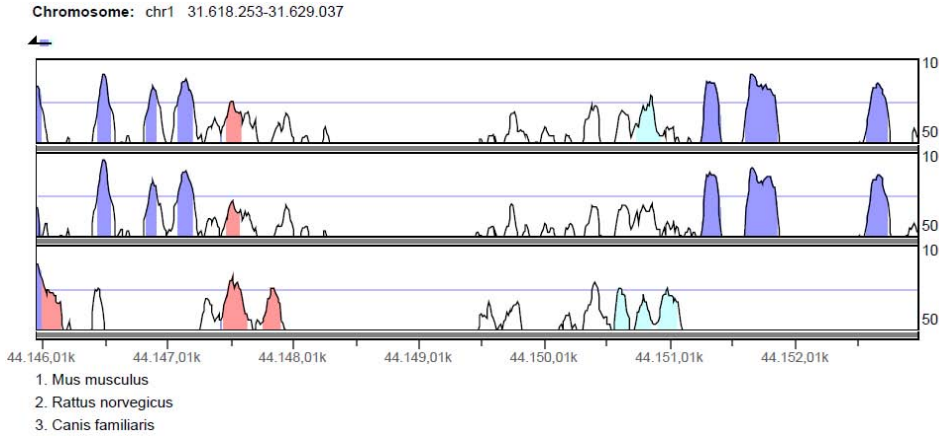
MYH6



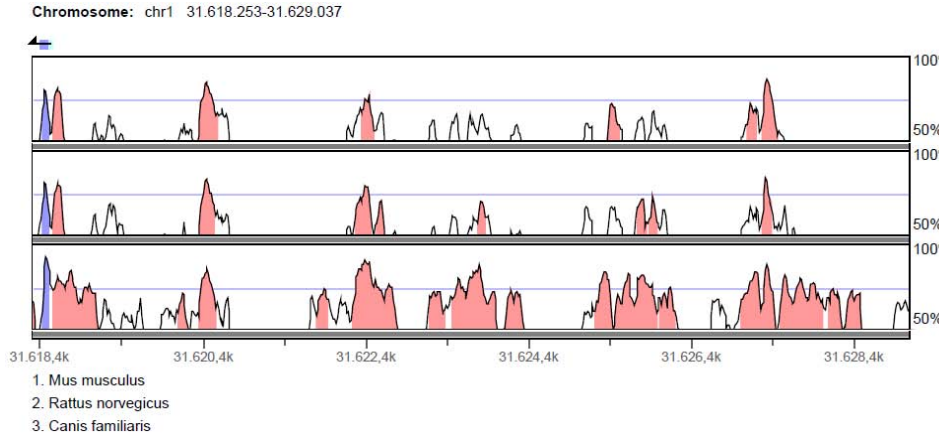
MYL3



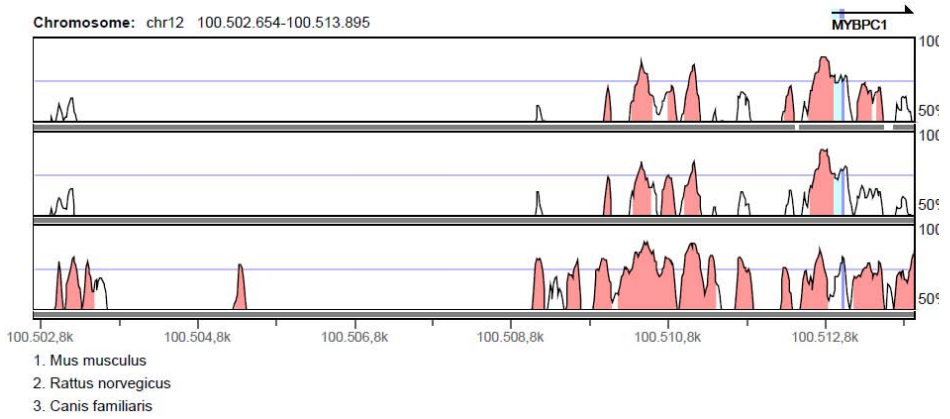
MYL7



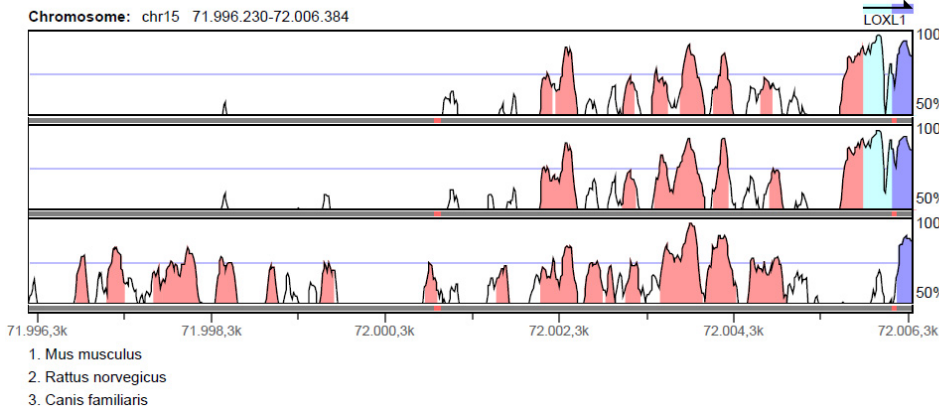
TCAP



MYBPC1



LOXL1



CHAPTER 4

ELECTRICAL STIMULATION OF PRIMARY NEONATAL RAT VENTRICULAR CARDIOMYOCYTES USING PACEMAKERS

Ruben S.R.M. Martherus; Volkert A. Zeilemaker; Torik A.Y. Ayoubi

Published

Biotechniques 2010 Jan; 48 (1) 65-67

Abstract

The study of gene regulation in cardiomyocytes requires a reliable *in vitro* model. However, monolayer cultures used for this purpose are typically not being exposed to electric stimulation although this has been shown to strongly affect cardiomyocyte gene expression. Based on pacemakers for clinical use we developed an easy to use and portable system that allows the user to perform electro-stimulation of cardiomyocyte cultures in standard tissue incubators without the need of bulky equipment. In addition we present here a refined protocol for culturing highly pure cardiomyocyte cultures with excellent contractile properties for a wide variety of applications.

Electrical stimulation of primary neonatal rat ventricular cardiomyocytes using pacemakers

Monolayer cultures of primary neonatal rat ventricular cardiomyocytes (NRVCMs) are a frequently employed model to study transcription regulation or cardiac-specific genes like atrial natriuretic factor (*Anf/ Nppa*) and alpha myosin heavy chain (*Myh6*).[1-3] However, many of such studies pass over on the importance of electric stimulation of cardiomyocytes although it has been found that prolonged electric-stimulation improves cardiomyocyte morphology and function and that electric pulsing directly increases transcript levels of cardiac-specific genes.[4, 5]

Currently, *in vitro* stimulation of cardiomyocytes or engineered cardiac tissue is done using cardiac stimulators like those of Nihon Kohden (Tokyo) or Ion Optix (Milton, MA. U.S.A) with parameters varying between 2 to 4 ms, 5 to 11 V/cm and 1 to 3 Hz.[6-8] Such apparatuses have disadvantages: they require to be plugged into the power grid and their large dimensions make transport to other work areas for instance for microscope analysis cumbersome. Furthermore, the machines normally have to be placed outside the incubators since they were often not designed for use in incubators with high humidity and temperatures. As a result structural modifications to the incubators may be required or connecting cables have to run through door-openings, potentially compromising sterility and culture conditions.

We hypothesized that employing cardiac pacemakers normally used for heart-patients would solve many of the aforementioned problems. These battery powered pacemakers are small and are completely closed so that the whole apparatus can be ethanol sterilized and placed inside incubators together with the cultures without the need for any modifications. Furthermore, the whole system together with cultures can easily be transported to any place and used there without interrupting the pulsing.

We cultured NRVCs in commercially available chamber-slides which can be used for a wide variety of purposes: live imaging, immunohistochemistry, reporter studies and a single chamber provides sufficient material for gene expression analysis. Reusable polycarbonate lids were designed to fit the chamber-slides and equipped with platinum electrodes running along opposite sides in each chamber. Small air channels in the lids enabled gas exchange (see figure 1). The electrode-lids were connected to human pacemakers, models Kappa 931 and 401 (Medtronic). The pacemakers were programmed using a Medtronic Vitatron 9790 programmer to give 7.5 V pulses (corresponding to 4.5 V/cm) with a duration of 1.5 ms at 1 Hz, placing it in the same range as previous publications and resembling the levels required for exciting normal ventricular tissue.[7-9] New pacemakers and programmers can be costly and may not be feasible for all researchers. However, significant numbers of devices are explanted for various reasons. These explanted pacemakers can be obtained and programmed with the help of cardiologists and may still be able to stimulate for several years.

The protocol employed for cardiomyocyte isolation and culture is a modified version of that described by Radisic et al.[9, 10] On day 1 hearts were obtained from 0-3 day old Lewis rats locally bred at the animal facility of Maastricht University. The use of animals for this study was evaluated and approved by the institute's animal ethical committee. On day 2 cardiomyocytes were pre-plated for 2 hours and subsequently plated at $>30,000$ cells/cm² on plastic (Permanox®, Lab-Tek) chamber-slides and incubated overnight in high serum medium supplemented with 10 µmol/L cytosine arabinoside (Ara-C) to further decrease the number of actively dividing (mostly non-myocyte) cells.[11, 12] The Permanox chamber-slides omitted the need for coatings since cardiomyocytes readily adhere to them. The culturing parameters were empirically determined through pilot experiments which revealed that the use of glass based chamber-slides with surface coatings like Histogrip™ (Invitrogen) and especially Matrigel™ (BD Biosciences) led to dramatic changes in cardiomyocyte cellular structure and function;

cardiomyocytes remained rounded, resulting in diminished intercellular connections and fewer contractions (data not shown).

At densities lower than 30,000 cells/cm² without electric stimulation the cells displayed no contractions, remained round and >50% of the cells died within 2 days of culture most likely because they were not able to form intercellular connections required for intercellular signaling. The resulting cultures consisted of >95% cardiomyocytes (purity of cultures can be observed from figure 2).

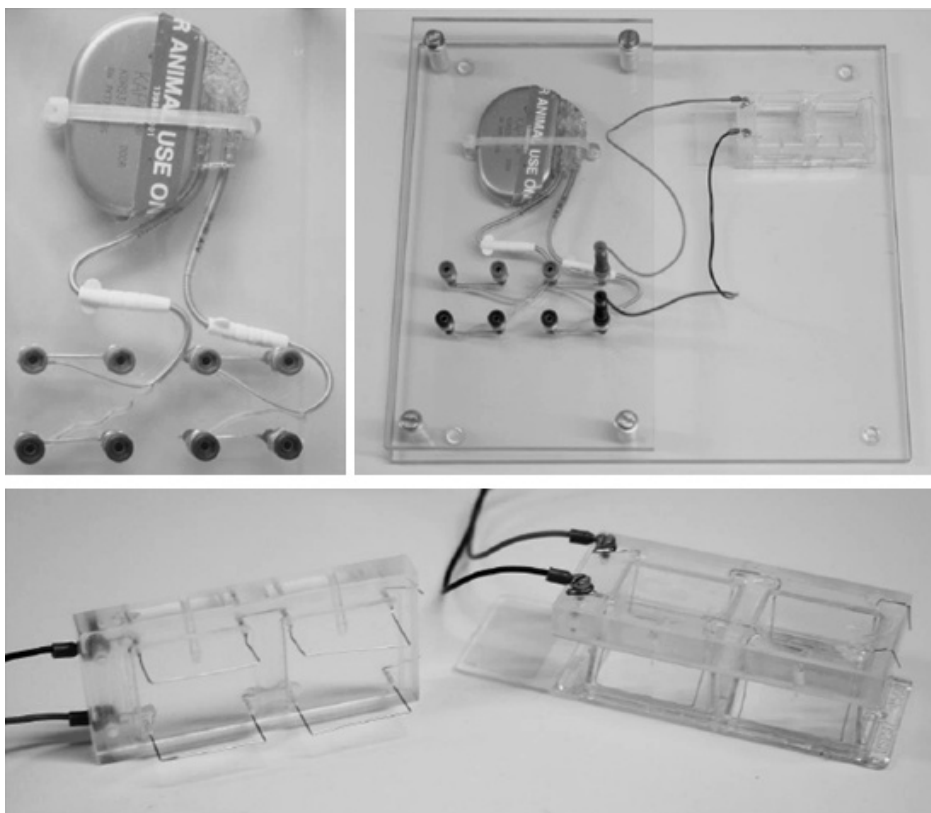


Figure 1. System for culturing electro-stimulated cardiomyocytes.

Top left panel: Human pacemaker Medtronic model KAPPA 931 fixed to polycarbonate platform. Ventricular and atrial leads are connected to two separate pairs of sockets enabling electro-stimulation of 2 series with different pulse characteristics. Top right panel: Overview of pulsing system stimulation; Bottom panel: Polycarbonate reusable electrode lids designed to fit commercial dual chamber slides.

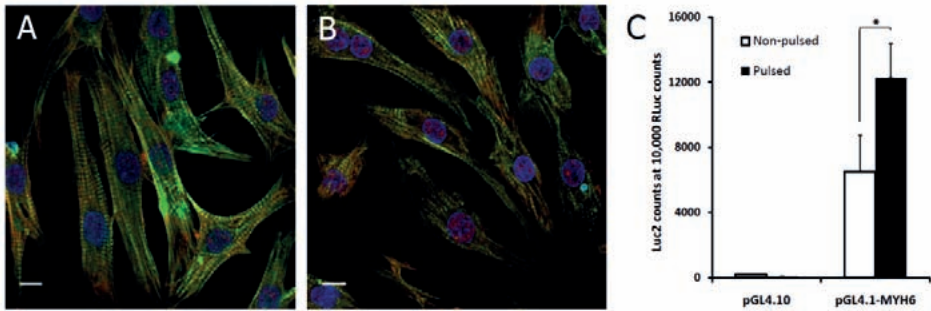


Figure 2. Gene expression in cultures of electro-stimulated cardiomyocytes.

Confocal microscopy images of cardiomyocytes after three days of electric stimulation (A) and without stimulation (B). Cells were cultured and fixed on chamber-slides and stained with mouse antibody for rat cardiac troponin-I and FITC-labeled goat-anti-mouse. Actin was stained with Alexa- 494 labeled phalloidin and nuclei with DAPI. Scale bar indicates 10 μ m. (C) Transcriptional activity of a 2kb human *MYH6* promoter fragment in cultures with and without stimulation 3 days post-transfection. Cells were transfected with equal molar amounts of reporter plasmids and 10 ng of pGL4.75 CMV-Renilla luciferase reporter vector.

On day 3 the medium was replaced with Ara-C-free high serum medium since Ara-C was shown to directly interfere with contractile activity. For transfection studies, cells can be transfected at this time. We obtained reproducible results for luciferase reporter constructs with the transfection reagent Fugene HD (Roche) using the manufacturers protocols when transfecting cells in high serum medium.

In our studies we employed transfection studies to examine the promoter sequences of cardiac genes. A *Myh6* promoter luciferase-reporter construct was transfected into cardiomyocytes which subsequently were electro-stimulated (figure 2). The behavior of the luciferase reporter construct is similar to that of the endogenous *Myh6*; the activation observed in figure 2 is believed to be electric pulsing-dependent.

Typical spontaneous contractions normally observed in cardiomyocytes became visible on day 4, also in absence of electric stimulation. From day 4 onwards, culture medium was refreshed every 24 hours with differentiation medium (DMEM, 2% horse serum, 1% HEPES, gentamycin 0.10 mmol/L, Fungizone 0.22 nmol/L) to avoid starvation of the cells. Serum concentration was shown to directly affect

formation of sarcomere structures and contractile activity. This was to be expected since serum levels directly modulate Serum Response Factor, a key cardiac transcription factor.[13, 14] Empirically we found normal differentiation and frequent spontaneous contractions in stimulated and non-stimulated cultures to be maintained at serum levels of 2% even after several days of culture. The best results for electro-stimulation were obtained when starting stimulation on day 4 together with the first addition of differentiation medium, thereby allowing for a recovery time of several days similar to previously described.[6]

In short, we present here a refined protocol for culturing NRVCs and a pacemaker-based pulsing system which inherent characteristics renders it possible to place the whole system inside any standard incubator, making it much more convenient in use than conventional equipment. The electro-stimulated NRVCs can be used for a many applications including reporter-assays, an interesting combination which to our knowledge remained un-investigated so far.

References

1. T.-g. Kim, J. Chen, J. Sadoshima, et al., Junonji Represses Atrial Natriuretic Factor Gene Expression by Inhibiting Transcriptional Activities of Cardiac Transcription Factors. *Mol Cell Biol.*, 24 (2004) 10151-10160.
2. K. Ojamaa, A.M. Samarel, and I. Klein, Identification of a Contractile-responsive Element in the Cardiac alpha-Myosin Heavy Chain Gene. *J Biol Chem.*, 270 (1995) 31276-31281.
3. T.-g. Kim, J. Jung, M.R. Mysliwiec, et al., Junonji represses alpha-cardiac myosin heavy chain expression via inhibiting MEF2 activity. *Biochem Biophys Res Commun.*, 329 (2005) 544-553.
4. P.M. McDonough and C.C. Glembotski, Induction of atrial natriuretic factor and myosin light chain-2 gene expression in cultured ventricular myocytes by electrical stimulation of contraction. *J Biol Chem.*, 267 (1992) 11665-11668.
5. E. Holt, P.K. Lunde, O.M. Sejersted, et al., Electrical stimulation of adult rat cardiomyocytes in culture improves contractile properties and is associated with altered calcium handling. *Basic Res Cardiol.*, 92 (1997) 289-298.
6. M. Radisic, H. Park, H. Shing, et al., Functional assembly of engineered myocardium by electrical stimulation of cardiac myocytes cultured on scaffolds. *PNAS*, 101 (2004) 18129-18134.
7. H. Yonemochi, S. Yasunaga, Y. Teshima, et al., Rapid Electrical Stimulation of Contraction Reduces the Density of beta-Adrenergic Receptors and

- Responsiveness of Cultured Neonatal Rat Cardiomyocytes. Possible Involvement of Microtubule Disassembly Secondary to Mechanical Stress. *Circulation*, 101 (2000) 2625-2630.
8. Rybkin, I.I., D.W. Markham, Z. Yan, et al., Conditional expression of SV40 T-antigen in mouse cardiomyocytes facilitates an inducible switch from proliferation to differentiation. *J Biol Chem.*, 278 (2003) 15927-34.
 9. M. Radisic, M. Euloth, L. Yang, et al., High-density seeding of myocyte cells for cardiac tissue engineering. *Biotechnol Bioeng.*, 82 (2003) 403-414.
 10. M. Radisic, A. Marsano, R. Maidhof, et al., Cardiac tissue engineering using perfusion bioreactor systems. *Nat. Protocols*, 3 (2008) 719-738.
 11. J.E. de Vries, M.M. Vork, T.H. Roemen, et al., Saturated but not mono-unsaturated fatty acids induce apoptotic cell death in neonatal rat ventricular myocytes. *J Lipid Res.*, 38 (1997) 1384-1394.
 12. P.J.H. Smeets, B.E.J. Teunissen, A. Planavila, et al., Inflammatory Pathways Are Activated during Cardiomyocyte Hypertrophy and Attenuated by Peroxisome Proliferator-activated Receptors PPAR-alpha and PPAR-delta. *J Biol Chem.*, 283 (2008) 29109-29118.
 13. R. Treisman, Identification of a protein-binding site that mediates transcriptional response of the c-fos gene to serum factors. *Cell*, 46 (1986) 567-574.
 14. J.M. Miano, X. Long, and K. Fujiwara, Serum response factor: master regulator of the actin cytoskeleton and contractile apparatus. *Am J Physiol Cell Physiol*, 292 (2007) C70-81.

CHAPTER 5

ELECTRICAL SIGNALS AFFECT THE CARDIOMYOCYTE TRANSCRIPTOME INDEPENDENTLY OF CONTRACTION

Ruben S.R.M. Martherus; Sabina J.V. Vanherle; Erika D.J. Timmer;
Volkert A. Zeijlemaker; Jos L. Broers; Hubert J.M. Smeets; Joep P.M.
Geraedts; Torik A.Y. Ayoubi

Published

Physiological Genomics. 2010 Nov 29;42A(4):283-289

Abstract

Cardiomyocytes *in vivo* are continuously subjected to electrical signals that evoke contractions and instigate drastic changes in the cells' morphology and function. Studies on how electrical stimulation affects the cardiac transcriptome have remained limited to a small number of heart-specific genes. Furthermore, these studies have ignored the interplay between the electrical excitation and the subsequent contractions. We carried out a genome-wide assessment of the effects of electrical signaling on gene expression, while distinguishing between the effects deriving from the electrical pulses themselves and the effects instigated by the evoked contractions. The changes in gene expression in primary cultures of neonatal ventricular cardiomyocytes from Lewis *Rattus norvegicus* were investigated using micro-arrays and RT-QPCR. A series of experiments was included in which the culture medium was supplemented with the contraction inhibitor blebbistatin to allow for electrical stimulation in the absence of contraction. Electrical stimulation was shown to directly enhance calcium handling and induce cardiomyocyte differentiation by arresting cell division, and activating key cardiac transcription factors as well as additional differentiation mechanisms such as Wnt-signaling. Several genes involved in metabolism were also directly activated by electrical stimulation. Furthermore, our data suggest that contraction exerts negative feedback on the transcription of various genes. Together, these observations indicate that intercellular electric currents between adjacent cardiomyocytes have an important role in cardiomyocyte development. They act at least partially through a pulse-specific gene-expression program that is activated independently from the evoked contractions.

Introduction

Electrical pulses are essential for a normal functioning heart. They regulate the rhythm of contractions of cardiomyocytes. Together with the induced contractions, they are also crucial for maintaining the normal cellular structure and functioning of the individual cardiomyocytes. Cultured adult cardiomyocytes display better contractile activity when the cells have been subjected to continuous electrical pulsing. Several studies have shown that the sarcomere

structure and intercellular structures such as gap junctions of cardiomyocytes are maintained when subjected to extracellular electrical stimuli, while in their absence, these structures are disrupted.(3, 13, 18, 20) A similar response to electrical pulsing was observed for skeletal myocytes. For these cells it was found that the changes in sarcomere organization were not caused by the higher expression levels of various main structural proteins.(9) However, the expression levels of the structural proteins myosin heavy chain 7 (*Myh7*) and myosin light chain 2 (*Myl2*) were both found to be elevated by electrical stimulation in cardiomyocytes.(18, 25) The increased expression of *Myl2* appeared to be mediated by voltage-gated calcium channels and calmodulin, indicating the direct involvement of calcium handling.(18)

The heart-specific hormone atrial natriuretic factor (ANF), encoded by the gene natriuretic peptide precursor A (*Nppa*), displayed a pulse-dependent increase in expression similar to *Myl2*.(18) However, ANF-levels also increase when cells are passively stretched, indicating that increased ANF expression could well be the result of the increased contractions induced by electrical stimulation and not by the increased electrical signaling itself.(14) In other studies it was found that mechanical stretching itself leads to changes in morphology, and differential RNA and protein synthesis. Furthermore, stretching was also found to affect intracellular calcium signaling in ways similar to electrical stimulation.(8, 15, 16, 25) A genome-wide analysis of mechanically induced changes in cardiomyocyte gene expression was undertaken by Frank and co-workers.(8) They postulated that the changes in gene expression that were observed in mechanically stretched cultures but not in phenylephrine-treated cultures or untreated controls were elicited by a stretch-specific gene program. A genome-wide study of the role of electrical signaling in the regulation of the cardiomyocyte transcriptome is currently lacking.

In this study we provide a genome-wide analysis of the electrical pulse-induced changes in gene expression, in neonatal rat ventricular cardiomyocytes, which are frequently used as a model in cardiac research. In order to distinguish between the responses to the pulse-induced calcium oscillations and subsequent contractions we employed the contraction inhibitor blebbistatin. We found that, independently from contraction, electrical signals directly affect the transcript levels of genes involved in a wide variety of processes including metabolism, cardiomyocyte differentiation and calcium handling, thus showing that

electrical signals by themselves play a crucial role in cardiomyocyte development and function.

Materials and methods

Cell culture

Hearts were obtained from 0–3-day-old male and female Lewis rats from the experimental animal facility of Maastricht University. The use of animals for this study was evaluated and approved by the Maastricht University animal ethical committee. Procedures were in accordance with institutional guidelines. Cardiomyocytes were isolated and cultured as described previously.⁽¹⁷⁾ After pre-plating, cells were plated in Permanox chamber slides (Lab-Tek) at 50,000 cells/cm². When applicable, cells were electrically stimulated for 3 days starting when the differentiation medium was first applied. For contraction inhibition the ‘differentiation’ medium was supplemented with 10 µmol/L blebbistatin (Sigma-Aldrich, St. Louis, MO, USA). For stimulation of contraction in the absence of electrical stimulation, phenylepinephrine (Sigma-Aldrich) was added to 50 µmol/L.

Immunohistochemistry

Cardiomyocytes grown on Permanox chamber slides (Lab-Tek) were washed briefly 2X with HBSS. Cells were fixed for 20 minutes in 3.7% paraformaldehyde in PBS (Invitrogen) pH 7.5. After washing three times briefly with PBS the cells were incubated for 45 minutes in blocking solution: 1% bovine serum albumin in PBS. Incubation with primary mouse monoclonal antibody to cardiac troponin I (Abcam, Cambridge, UK) was carried out in 1% BSA and 0.1% Tween-20 in PBS for 45 minutes. After three washes with 0.05% Tween-20 in PBS, samples were incubated for 30 minutes with the FITC-labeled goat-anti-mouse secondary antibody (Southern Biotech, Birmingham, AL, USA) and Phalloidin-alexa 595 (Invitrogen, Paisley, UK) in PBS containing 1% BSA and 0.1% Tween-20. Samples were then washed three times with 0.05% Tween-20 in PBS, twice with PBS, and mounted in 90% glycerol, containing 20 mM Tris-HCl pH 8.0, 0.2% NaN₃, and 2% 1,4-diazobicyclo-(2,2,2)-octane (DABCO; Merck, Darmstadt, Germany). Nuclei were counterstained using 0.5 µg/ml diamidino-2-phenylindole (DAPI;

Sigma-Aldrich). Images were obtained using a Leica TCS SPE confocal laser scanning microscope system with a DMI 4000B inverted microscope. Cell diameter measurements were performed using ImageJ (NIH, USA).

Live imaging

Films were made using a Leica DMRI BE microscope equipped with a Hamamatsu digital camera (type C4742-95). The individual images, of which the films were composed, were recorded using a 10X magnification and an exposure time of 0.1 seconds. The final movies displayed the cells in real time. Open Lab version 2.2.5 (Improvision, Waltham, MA, USA) software was used. Using a Thermoplate (Tokai HIT, Shizuoka-ken, Japan), the mean temperature was maintained at 37 °C.

RT-QPCR

RNA was isolated using Trizol reagent (Invitrogen, Paisley, UK). First strand cDNA was made using M-MLV reverse transcriptase (Finnzymes, Espoo, Finland) starting from 1.0 µg of total RNA. QPCR was performed using a qPCR SYBR-Green master mix (Eurogentec, Liège, Belgium) in 10-µl reactions on an ABI7900 (Applied Biosystems, Foster City, CA, USA). In all experiments peptidylprolyl isomerase-A (*Ppia*) was used as the reference gene. The oligonucleotide sequences are provided in the supplementary data.

Expression arrays

For each array, two separate samples from one series of either pulsed or non-pulsed cells were pooled. RNA quantity and quality were checked using Bioanalyzer RNA nano chips (Agilent, Santa Clara, CA, USA). One µg total RNA was taken from pooled RNA samples and processed using an Affymetrix One-Cycle cDNA synthesis kit (Affymetrix, Santa Clara, CA, USA). Samples were applied to a GeneChip Rat Expression Set 230 (Affymetrix) and processed according to the manufacturer's instructions. CEL-files were transformed using the PLIER-algorithm. Analyses for differentially expressed genes were performed using BRB-Array Tools version 3.7.0 (<http://linus.nci.nih.gov/BRB-ArrayTools.html>). In short, the array data were normalized using the housekeeping gene normalization option in

BRB-Array tools. For housekeeping genes we selected the rat orthologs of the housekeeping genes of the Affymetrix HG-U133 plus 2.0 arrays (the list of housekeeper probe sets is provided in the supplementary data). The resulting data were analyzed for differential expression of genes through class comparison of three separate series of experiments with arrays paired by experimental series (by batch of cardiomyocytes). To identify the affected biological processes, genes that were significantly differentially expressed between groups ($p < 0.05$) were classified using DAVID (Database for Annotation, Visualization and Integrated Discovery).(7) We listed the clusters with an enrichment score ($-\log_{10}p$) corresponding to $p < 0.05$. The microarray data discussed in this publication were deposited in NCBI's Gene Expression Omnibus and are accessible through GEO Series accession number GSE15856: [\(1\)](http://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?acc=GSE15856)

RESULTS

Microscopic analysis of cell morphology and contractile activity.

In our cell culture procedure the cardiomyocytes first underwent partial de-differentiation during which sarcomeric structures diminished and cells decreased in size. We found this to be the result of the low temperature (4 °C) at which the overnight incubation was performed (data not shown). In the initial recovery phase, upon plating, the cells increased in size again, formed cell-cell connections and eventually displayed the characteristic frequent spontaneous contractions. We performed a more detailed analysis of the influence of electrical stimulation on cardiomyocyte cell morphology using immunohistochemistry and confocal microscopy (figure 1). Cultures were fixed and stained after 3 days of culture in low serum differentiation medium with or without electrical stimulation. In addition, we included a series of non-stimulated cells that were supplemented with epinephrine.

A difference in the overall shape of the cells was observed between electrically pulsed and non-pulsed cells. Non-pulsed cells displayed a somewhat fibroblast-like elliptical shape while the electrically pulsed cells were more 'straight/rectangular' over their whole length, similar to the cardiomyocytes in the native heart. The epinephrine-stimulated cells showed the typical appearance of hypertrophy, characterized by the increased width of the cells (quantified and depicted in Figure 1J).

Striation patterns originating from stained sarcomere protein structures were visible in all samples, although the epinephrine-treated cells showed a more diffuse pattern when compared to the electrically stimulated cells. Separate myofibrils remained clearly visible in the non-pulsed cardiomyocytes. In electrically stimulated cardiomyocytes and those treated with epinephrine the myofibrils were packed closely against each other over the whole width of the cell with no visible spaces in between.

The contractile activity of living cardiomyocytes in pulsed and non-pulsed cultures was visualized and recorded using light microscopy. Movies displaying the contractions in real time are available as on-line supplementary data. As shown in these movies, the cardiomyocytes that were subjected to electrical pulsing displayed more constant and frequent contraction when compared to non-stimulated cells. The rapid synchronized contractions occurred more frequently than those at the 1 Hz at which the pacemaker was programmed. We believe that this was the result of spontaneous contraction of the cardiomyocytes themselves, because the contractile activity continued when stimulation was temporarily halted (results not shown).

Pulsing induced changes in gene expression.

We evaluated the effect of extracellular electrical stimulation on the expression of gap junction protein alpha 1/connexin-43 (*Gja1/Cx43*), myosin heavy chain 6 (*Myh6*), myosin heavy chain 7 (*Myh7*) and *Nppa* in cardiomyocytes at different cell densities using RT-QPCR analysis. In these preliminary experiments we observed that some of the pulse-induced changes were diminished at higher cell densities (>100,000 cells/cm², see Figure 2). Microscopic inspection of living cells at those higher cell densities revealed no difference in contractile activity, and both pulsed and non-pulsed cultures displayed contractions at frequencies higher than 1 Hz. We suspected that at higher cell densities depolarization waves originating from the spontaneous contractions overpowered the pulses supplied by the pulsing system. In order to detect as much pulse-related change as possible we used a seeding cell density of 50,000 cells/cm² in all subsequent experiments. The frequencies at which the employed pacemakers can pulse are limited to those found in the human *in vivo* situation. Therefore we kept the pulsing frequency at 1 Hz although the intrinsic rhythm of neonatal heart-beat frequency is higher than 1 Hz. In preliminary experiments we

found that significant changes in cell morphology and gene expression were observed even at this frequency.

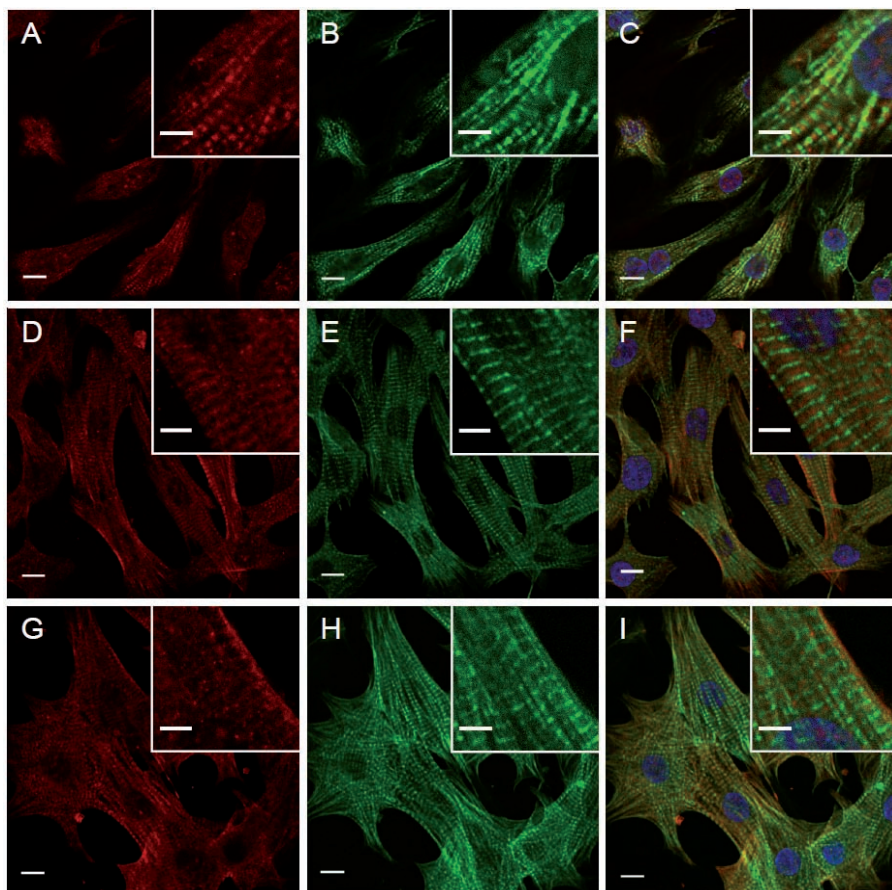


Figure 1. Effect of extracellular applied electrical pulsing on cardiomyocyte morphology.

Cardiomyocyte cultures were fixed and stained after 3 days of culture in low-serum medium, and were non-pulsed (A-C), pulsed (D-F) or non-pulsed but supplemented with phenylepinephrine (G-I). The structural proteins actin (A,D,G) and the cardiac-specific troponin-I (B,E,H) were detected using phalloidin and *TNNI3*-specific antibody staining. An overlay shows DAPI-stained nuclei (C,F,I). Scale bars indicate a length of 10 μ m in the overview pictures and 4 μ m in the enlarged views.

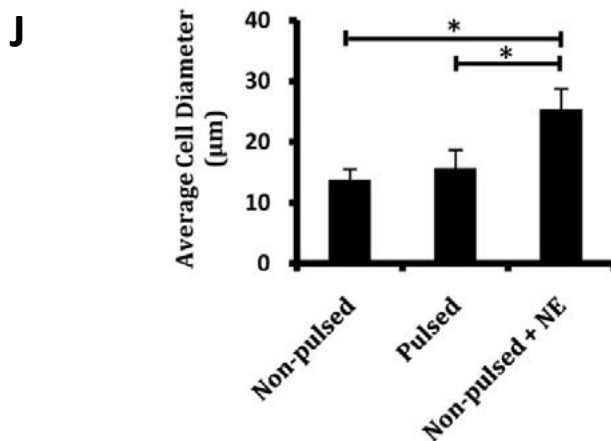


Figure 1. Continued from previous page

To compare cell diameter, an indicator for hypertrophy, cellular diameters were measured at the site of the nuclei and average values were calculated (J). * Indicates a significant difference ($p < 0.006$) according to the t-test.

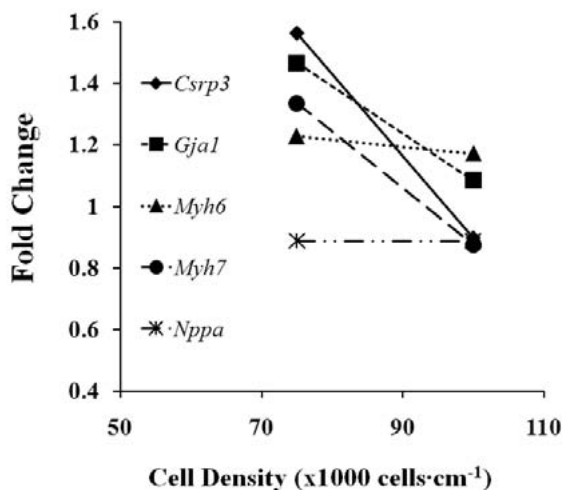


Figure 2. Susceptibility to electrical pulsing at different cell densities.

Samples from cultured extracellular electrically stimulated (3 days of stimulation) and non-stimulated cardiomyocytes were subjected to RT-QPCR analysis for the genes, *Csrp3*, *Gja1/Cx43*, *Myh6*, *Myh7* and *Nppa*. Changes in gene expression are expressed as the fold change of stimulated cultures over non-stimulated cultures of the same cell density.

To obtain an integral overview of differentially expressed genes as a result of electrical stimulation we performed gene-expression microarray analysis. Differentially expressed genes were identified by a pair-wise comparative analysis using BRB-Array Tools. Probe sets with $p < 0.05$ were considered to be significantly differentially expressed. Out of the 2474 differentially expressed probe sets, the vast majority (1948) was up-regulated (see supplementary data). Differentially expressed genes were analyzed through gene ontology to identify the biological processes, cellular compartments and molecular functions in which the genes were enriched (Table 1).

Table 1: Gene ontology clusters enriched with genes differentially expressed in electrically stimulated cardiomyocytes.

Cluster	Sub-cluster	Up-regulated	Down-regulated
Biological process	Apoptosis		1
	Cyclase activity positive regulation	4	
	Differentiation / morphogenesis	6	
	Heart rate, positive regulation	3	
	Transcription regulation	1	
	Transcription, negative regulation	5	
	Transcription, positive regulation	7	
	Vasculature development	2	2
Cellular compartment	Mitochondrion	2	
	Myosin filament	3	
	Sarcomere	1	
Molecular function	Ion channel activity	1	

Genes differentially expressed ($p < 0.05$) by electrical pulsing were functionally annotated through clustering based on gene ontology terms using the DAVID program. Annotated clusters are ranked alphabetically within their category. Numbers refer to the ranking of the enrichment score, 1 being the most enriched. The DAVID program does not supply cluster names; therefore we supplied our own cluster names reflecting the elements of the clusters.

Hallmarks among the biological processes affected by electrical pulsing were the up-regulated ‘differentiation/ morphogenesis’ and ‘positive regulation of heart rate’. Apoptosis was found to be down-regulated by electrical pulsing. Furthermore, we observed the increased expression of genes encoding components of the cardiomyocytes’ sarcomeres and ion-channels. In Table 2 we provide the gene-expression data for a selection of genes within these groups. For a number of genes we verified gene expression using RT-QPCR (Figure 3). Although the changes observed in the RT-QPCR were much stronger, the overall changes followed the same pattern.

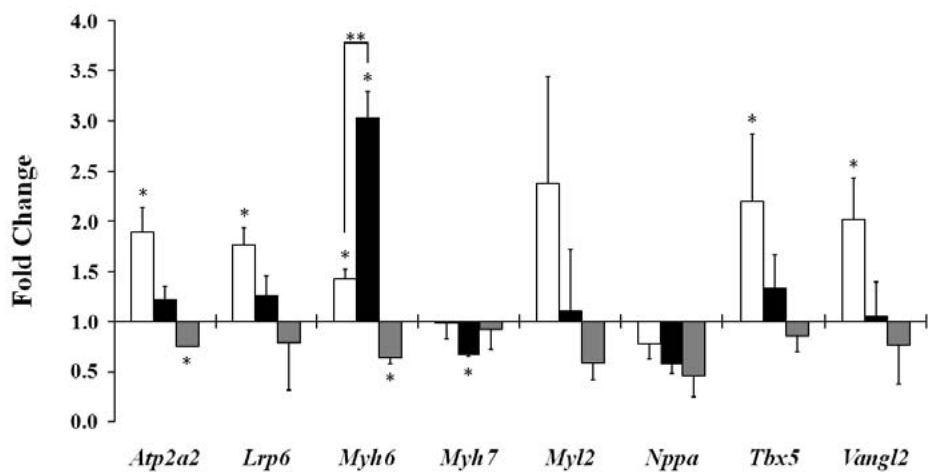


Figure 3. Electrical stimulation affects transcription independently from contraction.

Graph depicts fold changes of gene expression as determined through RT-QPCR analysis on samples from cultured extracellular electrically stimulated (3 days of stimulation) and non-stimulated cardiomyocytes. White bars: comparing electrically stimulated to non-stimulated. Black bars: comparing blebbistatin-supplemented electrically stimulated to normal stimulated. Grey bars: comparing blebbistatin-supplemented non-stimulated to normal non-stimulated.

* indicates significant ($p < 0.05$) differential expression. ** indicates a significant difference ($p < 0.05$) between the two series.

Table 2: Overview of selected differentially expressed genes in cardiomyocytes as determined through expression microarray analysis

Cluster	Probe set	Refseq ID	Gene symbol	Fold change Pulsed / Non-pulsed	Fold change Pulsed + blebbistatin/ Pulsed
Ion channels (and related functions)					
	1398862_at	J04024	Atp2a2	1.62*	1.17
	1370452_at	M59786	Cacna1c	1.37*	1.05
	1377499_at_at	NM_181369	Hrc	1.76*	1.67*
	1368005_at	NM_013138	Itpr3	0.73*	0.88
	1368560_at	NM_017297	Kcnj5	1.76*	0.97
	1370198_at	AJ243304	Trdn	1.75*	1.68*
Sarcomere structure					
	1387049_at	B1296041	Myh6	1.62	1.60*
	1386993_at	NM_017240	Myh7	1.26	0.75*
	1371315_at	AA891242	MyI7	1.41*	1.70*
	1372195_at	NM_001037351	Tnnc2	0.77*	1.09
	1390061_at	B1289643	Tnnt2	1.67*	0.79

Table 2: (continued)

Differentiation / morphogenesis				
	1376725_at	BM391816	lrp6	1.35*
	1373410_at	BE104219	Mef2c	1.25*
	1385125_at	NM_182667	Myocd	1.58*
	1377139_at	BE110811	Tbx5	1.55*
	1387750_at	NM_053530	Twist1	0.80*
	1392667_at	AI070793	Vangl2	1.30*
Metabolism				
	1380557_at	NM_012782	Bckdha	1.53*
	1367742_at	NM_013200	Cpt1b	1.53*
	1382986_at	NM_017274	Gpam	1.48*
	1370237_at	NM_057186	Hadh	1.50*

A selection of genes with differential expression as calculated using BRB-array-tools ($p < 0.05$, indicated by *) are listed according to the processes or structures in which they are involved. The fold changes resulting from electrical pulsing of standard cultures and from the addition of the contraction inhibitor blebbistatin to electrically pulsed cultures are shown. Complete datasets are provided in the supplementary data.

Blebbistatin supplementation reveals a stretch-independent gene expression program.

The observed changes in gene expression induced by electrical stimulation were to some extent the result of increased contractile activity. In order to identify which of the changes in the gene expression program arose specifically by pulsing, independently from contraction, we supplemented electrically pulsed cardiomyocyte cultures with the contraction-blocking agent blebbistatin. In Table 2 the ratios of fold changes determined via microarray analysis are given as indicators of the changes in expression between matching pairs of experiments. For many genes, supplementation with blebbistatin did not lead to a significant change in expression levels, indicating that the increase observed in the pulsed cultures (compared to non-pulsed) resulted only from electrical stimulation. A few genes, including triadin (*Trdn*), *Myh6*, myosin light chain 7 (*Myl7*), carnitine palmitoyltransferase 1b (*Cpt1b*) and hydroxyacyl-Coenzyme A dehydrogenase (*Hadh*) displayed an even further increase in expression. Through RT-QPCR, we confirmed this for *Myh6* (Figure 3). When considering the effects of pulsing and blebbistatin supplementation of the pulsed cells, the cumulative changes in gene expression of some of the genes indicated in Table 2 differed by more than three-fold compared to non-pulsed cells.

To investigate the possible effect of blebbistatin addition itself on gene expression we compared the expression levels of selected genes in non-stimulated cultures with and without blebbistatin supplementation (Figure 4). In the absence of external stimulation, the addition of blebbistatin led to a decrease in the expression of *Myh6* and *Atp2a2*. Although not reaching significance, the cardiac-specific genes *Myl2*, *Nppa* and *Tbx5* also displayed an overall trend in down-regulation.

Discussion

Using a pulsing system based on human pacemakers, we carried out a genome-wide assessment of the effect of electrical stimulation on the transcriptome of cultured primary neonatal rat ventricular cardiomyocytes.(17) Immunohistochemistry and imaging of living cells in culture showed that electrical stimulation appeared to enhance (re-)differentiation. This proposed increase in differentiation was confirmed by gene-expression analyses. We found that electrical

stimulation led to the increased expression of cardiac-specific genes such as *Myh6*, *Gja1* and the cardiac-specific isoform of troponin-T (*Tnnt2*). At the same time, myosin light chain 1(*Myl1*) and troponin-C 2 (*Tnnc2*) were down-regulated. Both these genes are known to have a more skeletal muscle-restricted expression profile. The cells thus deviated more from a common striated muscle phenotype. With regards ion channels, we observed increased expression of cardiac-specific transcripts, amongst them sarcoplasmic calcium channel *Atp2a2* and the cardiac ventricle L-type Ca^{2+} channel isoform *Cacna1c*.(11)

Activation of the cardiac-specific transcription program was further demonstrated by the differential expression of several key cardiac-enriched transcription factors. Increased expression was observed for T-box 5 (*Tbx5*), myocardin (*Myocd*) and myocyte enhancing factor C (*Mef2c*). All are involved in cardiogenesis and regulation of the transcription of heart-specific genes in later stages of life.(4, 22) Meanwhile, expression of the twist homolog *Twist1*, a negative regulator of cardiac-specific gene transcription, was decreased.(10) The increased differentiation was also apparent from the increased expression of some of the more ubiquitously expressed genes: increased levels of the cyclin-dependent kinases *Cdkn1c* and *Cdkn2b*, both of which augment the arrest of cell division, showed that the cells exit from a proliferative state and attain a more differentiated state. Activation of Van Gogh-like 2 (*Vangl2*) and low density lipoprotein 6 (*Lrp6*) indicated activation of Wnt-frizzled signaling.(19, 23)

The accelerated differentiation due to the application of externally generated electrical pulses was only observed in cardiomyocytes cultured at low densities ($\leq 50,000$ cells/cm²). It seems likely that in cardiomyocyte cultures at lower cell densities the extracellular electrical pulses functionally compensate for the lack of electrical stimuli normally obtained from the surrounding cells. Once the cultured cardiomyocytes develop further and form more intercellular connections, the cells can take over some of the pacemaker pulsing activity through their spontaneous contraction and provoke contractions in neighboring cells. This in itself also demonstrates that electrical signals transmitted by the cells themselves are important in cardiomyocyte development. Furthermore, these observations concur with the accelerated differentiation of mesenchymal stem cells to cardiomyocytes at higher cell densities, independently of extra electrical stimuli, and the improvement in cardiomyocyte function at higher cell densities.(6, 24)

At higher cell densities, the expression levels of *Myh6* were still increased as a result of electrical pulsing, suggesting a similar ability in

vivo. In a recent study on human heart failure patients, expression of *MYH6*, together with that of *ATP2A2*, was indeed increased in hearts receiving electrical stimulation.(5) In addition, our data showed that electrical pulsing leads to the down-regulation of genes involved in apoptosis, thus providing an additional explanation for the observed beneficial effect that electrical stimulation has on the failing heart.

Contrary to previous publications, we did not observe increased expression of *Nppa* in electrically stimulated cardiomyocytes.(18) Besides the unchanged expression of *Nppa*, the expression of the hypertrophy mediators *Mapk1* (ERK2), *Mapk3* *Map2k5* (ERK1) *Map2k5* (MEK5) and v-akt murine thymoma viral oncogene homolog 1 (*Akt1*) also remained unchanged by electrical stimulation.(2) In addition to this we found the expression of *Trdn* to be increased by pulsing while in PE-stimulated cells, a model for hypertrophy, a significant decrease has previously been observed.(8) Besides supporting the proposed absence of hypertrophy this shows that epinephrine and electrical stimulation activate very distinct mechanisms in gene expression regulation.

To distinguish between transcriptional responses to electrical currents and stretching, the contraction inhibitor blebbistatin was employed. For genes in which expression in pulsed cells was not altered by blebbistatin supplementation, it is likely that their observed up-regulation by electrical stimulation resulted from the electrical signaling and not from the subsequent contractions. *Tbx5* and *Mef2C*, encoding two key cardiac transcription factors, were among these genes, as well as glycerol-3-phosphate acyltransferase (*Gpam*) and alpha-branched chain ketoacid dehydrogenase (*Bckdha*), both of which are involved in β -oxidation.

Blebbistatin supplementation of electrically stimulated cultures led to an even further increase in the expression of several genes, e.g. *Myh6*, *Trdn*, *Cpt1b* and *Hadh* when compared to normal pulsed cultures. The latter two are also involved in β -oxidation and *Hadh* was previously described to be up-regulated during increased muscle activity.(21) This suggests that pulsing, independently from contraction, affects the expression of genes involved in physiological processes.

The down-regulation of *Myh6* in control experiments, with blebbistatin supplementation of non-stimulated cells, suggests that the increase in *Myh6* expression in the electrically stimulated cultures was not caused by a side effect of blebbistatin, but by the occurrence of a negative transcription regulatory mechanism exerted by stretching. Indeed, in the data from the stretch study by Frank et al., *Myh6* displayed a clear trend towards being down-regulated by stretching.(8)

A likely explanation for the down-regulation of *Myh6* and *Atp2a2* by blebbistatin in non-stimulated cultures is that overall cardiomyocyte development and function were seriously hampered when contraction was blocked in the absence of external electrical stimuli. It thereby affected the cardiomyocyte transcriptome as a whole.

The increase in expression of *Atp2a2*, *Cacna1c*, *Trdn* and *Hrc* and the down-regulation of the Ca^{2+} oscillation antagonist type 3 inositol 1,4,5-triphosphate receptor (*Itpr3*) suggest that calcium handling is activated in a similar way to that described in earlier publications.(12, 13) In conjunction with this we found that the activation of several genes persisted in blebbistatin-supplemented cultures. This affirmed that calcium cycling continued in the absence of contractions.

In addition to the identification of pulse-specific genes, the blocking of contraction by blebbistatin also revealed that several genes are to some extent activated by stretching. One such stretch-dependent gene is *Myh7*, although it was not identified as such in the mechanical stretch study by Frank et al.(8) Differences in cell density and the spontaneous contractions of cardiomyocytes are a likely cause of this. In their mechanical stretch study, Frank et al. used a cell density of 175,000/cm² while in our experiments cell density was typically 50,000/cm². At higher cell densities, cells can more easily excite each other electrically and induce deformation of the surrounding cells, thereby triggering the stretch-related gene expression program. Although not reaching significance, *Nppa* also displayed a slight trend for being down-regulated by blebbistatin supplementation. This supports the findings of Kinunnen et al., who reported *Nppa* to be activated by stretching.(14)

However, the use of neonatal myocytes at the low densities in this study brings with it some limitations for making comparisons to the *in vivo* situation of the (adult) heart. Pulse-induced changes in gene expression would have to be studied at higher cell densities (>100,000 cells/cm²) to obtain more representative information on the role of electrical signaling *in vivo*.

Nevertheless, the processes activated by electrical stimulation in the absence of contractions are also likely candidates for being induced by pulsing *in vivo*. Based upon our findings and the previously described activation of for example *ATP2A2* by electrical pulsing *in vivo*, we believe that pulse-dependent activation of calcium handling, as observed in our study, also occurs *in vivo*, and with it the fundamental effects that it exerts on the cardiomyocyte transcriptome.(5) Electrical signaling, by itself, is therefore also likely to provide an additional mechanism along

with contraction, through which the cardiomyocyte transcriptome reacts to alterations in heart activity.

We thus showed that aside from initiating contractions, electrical signals activate a unique and contraction-independent gene expression program. This program is important in cardiomyocyte development and function and acts through mechanisms that differ from those activated by mechanical stretching or epinephrine stimulation.

Supplementary data

Supplemental file 1 - Expanded-materials-and-methods.pdf

Supplemental file 2 - movie_pulsed.mov

Supplemental file 3 - movie_non-pulsed.mov

Supplemental file 5 - Micro-array-analysis-data.xls

Supplementary data is available on-line via:

<http://physiolgenomics.physiology.org/content/42A/4/283.full>

Grants

The work described here is part of a project funded by the Transnational University Limburg.

References

1. Barrett T, Troup DB, Wilhite SE, Ledoux P, Rudnev D, Evangelista C, Kim IF, Soboleva A, Tomashevsky M, Edgar R. NCBI GEO: mining tens of millions of expression profiles--database and tools update. *Nucleic Acids Res.* 35: D760-765, 2007.
2. Barry SP, Davidson SM, Townsend PA. Molecular regulation of cardiac hypertrophy. *The International Journal of Biochemistry & Cell Biology* 40: 2023-2039, 2008.
3. Berger HJ, Prasad SK, Davidoff AJ, Pimental D, Ellingsen O, Marsh JD, Smith TW, Kelly RA. Continual electric field stimulation preserves contractile function of adult ventricular myocytes in primary culture. *Am J Physiol Heart Circ Physiol.* 266: H341-349, 1994.
4. Bi W, Drake CJ, Schwarz JJ. The transcription factor MEF2C-Null mouse exhibits complex vascular malformations and reduced cardiac expression of angiotensin 1 and VEGF. *Developmental Biology* 211: 255-267, 1999.

5. Butter C, Rastogi S, Minden H-H, Meyhöfer J, Burkhoff D, Sabbah HN. Cardiac contractility modulation electrical signals improve myocardial gene expression in patients with heart failure. *J Am Coll Cardiol.* 51: 1784-1789, 2008.
6. Clark WA, Decker ML, Behnke-Barclay M, Janes DM, Decker RS. Cell contact as an independent factor modulating cardiac myocyte hypertrophy and survival in long-term primary culture. *J Mol Cell Cardiol.* 30: 139-155, 1998.
7. Dennis G, Sherman B, Hosack D, Yang J, Gao W, Lane HC, Lempicki R. DAVID: Database for Annotation, Visualization, and Integrated Discovery. *Genome Biol.* 4: P3, 2003.
8. Frank D, Kuhn C, Brors B, Hanselmann C, Ludde M, Katus HA, Frey N. Gene expression pattern in biomechanically stretched cardiomyocytes: evidence for a stretch-specific gene program. *Hypertension* 51: 309-318, 2008.
9. Fujita H, Nedachi T, Kanzaki M. Accelerated de novo sarcomere assembly by electric pulse stimulation in C2C12 myotubes. *Exp Cell Res.* 313: 1853-1865, 2007.
10. Hamamori Y, Wu HY, Sartorelli V, Kedes L. The basic domain of myogenic basic helix-loop-helix (bHLH) proteins is the novel target for direct inhibition by another bHLH protein, Twist. *Mol Cell Biol.* 17: 6563-6573, 1997.
11. Hatano S, Yamashita T, Sekiguchi A, Iwasaki Y, Nakazawa K, Sagara K, Iinuma H, Aizawa T, Fu LT. Molecular and electrophysiological differences in the L-type Ca^{2+} channel of the atrium and ventricle of rat hearts. *Circ J.* 70: 610-614, 2006.
12. Hattori M, Suzuki AZ, Higo T, Miyauchi H, Michikawa T, Nakamura T, Inoue T, Mikoshiba K. Distinct roles of inositol 1,4,5-trisphosphate receptor types 1 and 3 in Ca^{2+} signaling. *J Biol Chem.* 279: 11967-11975, 2004.
13. Holt E, Lunde PK, Sejersted OM, Christensen G. Electrical stimulation of adult rat cardiomyocytes in culture improves contractile properties and is associated with altered calcium handling. *Basic Res Cardiol.* 92: 289-298, 1997.
14. Kinnunen P, Vuolteenaho O, Uusimaa P, Ruskoaho H. Passive mechanical stretch releases atrial natriuretic peptide from rat ventricular myocardium. *Circ Res.* 70: 1244-1253, 1992.
15. Lammerding J, Kamm RD, Lee RT. Mechanotransduction in cardiac myocytes. *Ann. N.Y. Acad. Sc.* 1015: 53-70, 2004.
16. Mann DL, Kent RL, Cooper GT. Load regulation of the properties of adult feline cardiocytes: growth induction by cellular deformation. *Circ Res.* 64: 1079-1090, 1989.
17. Martherus RSRM, Zeijlemaker VA, Ayoubi TAY. Electrical stimulation of primary neonatal rat ventricular cardiomyocytes using pacemakers. *Biotechniques* 48: 305-307, 2010.

18. McDonough PM, Glembotski CC. Induction of atrial natriuretic factor and myosin light chain-2 gene expression in cultured ventricular myocytes by electrical stimulation of contraction. *J Biol Chem.* 267: 11665-11668, 1992.
19. Piao S, Lee SH, Kim H, Yum S, Stamos JL, Xu Y, Lee SJ, Lee J, Oh S, Han JK, Park BJ, Weis WI, Ha NC. Direct inhibition of GSK3 β by the phosphorylated cytoplasmic domain of LRP6 in Wnt/ β -catenin signaling. *PLoS ONE* 3: e4046, 2008.
20. Radisic M, Park H, Gerecht S, Cannizzaro C, Langer R, Vunjak-Novakovic G. Biomimetic approach to cardiac tissue engineering. *Philos Trans R Soc Lond B Biol Sci.* 362: 1357-1368, 2007.
21. Stuewe SR, Gwartz PA, Agarwal N, Mallet RT. Exercise training enhances glycolytic and oxidative enzymes in canine ventricular myocardium. *J Mol Cell Cardiol.* 32: 903-913, 2000.
22. Takeuchi JK, Ohgi M, Koshiba-Takeuchi K, Shiratori H, Sakaki I, Ogura K, Saijoh Y, Ogura T. Tbx5 specifies the left/right ventricles and ventricular septum position during cardiogenesis. *Development* 130: 5953-5964, 2003.
23. Taylor J, Abramova N, Charlton J, Adler PN. Van Gogh: a new *Drosophila* tissue polarity gene. *Genetics* 150: 199-210, 1998.
24. Wang T, Xu Z, Jiang W, Ma A. Cell-to-cell contact induces mesenchymal stem cell to differentiate into cardiomyocyte and smooth muscle cell. *Int J Cardiol.* 109: 74-81, 2006.
25. Xia Y, Buja LM, Scarpulla RC, McMillin JB. Electrical stimulation of neonatal cardiomyocytes results in the sequential activation of nuclear genes governing mitochondrial proliferation and differentiation. *Proc Natl Acad Sci USA* 94: 11399-11404, 1997.

CHAPTER 6

IDENTIFICATION OF PROMOTER ELEMENTS PREREQUISITE FOR CARDIAC SPECIFICITY AND THE CARDIAC RESPONSE TO ELECTRIC STIMULI

Ruben S.R.M. Martherus; Wouter Van Delm; Erika D.J. Timmer;
Sabina J.V. VanHerle; Yves Moreau; Hubert J.M. Smeets; Torik A.Y.
Ayoubi

Manuscript in preparation

Abstract

Of only a few cardiac genes the promoter architecture and associated transcriptional regulation have been investigated into great detail. However, for most of the cardiac genes there still remain many questions on the fundamentals of their transcriptional regulation. We screened the immediate 5' flanking promoter regions of cardiac genes for regulatory motifs and modules involved in obtaining cardiac-specific expression, and in addition for elements involved in the transcriptional response to electric stimuli. Similar to others, we found the elements for serum response factor (SRF), myocyte enhancer factor 2 (MEF2) and SP1 to be involved in obtaining cardiac specificity. Our screening for elements mediating the transcriptional response to electric stimulation revealed an enrichment of binding sites for transcription factors known to act in hypertrophy, TBX5, MEF2, GATA1, and LEF1. Through experimental validation we show that in the cysteine and glycine-rich protein 3 (*CSRP3*) promoter TCF7L2 binds to a predicted LEF1 site. While β -catenin signaling has been previously found to be involved in activity induced changes in gene-expression, we show that the binding sites involved, are enriched in the promoters of genes responding to pulsing in the absence of contraction.

Introduction

With the human genome sequence known, one of the challenges remaining is to determine which of the billions of base pairs in the DNA are involved in regulating the transcription of the more than 20 thousand genes encoded within the human genome. To circumvent the laborious research on the transcriptional regulation of individual genes, bioinformatics analyses are performed on sets of genes that respond in parallel to specific stimuli or display similar spatial and or temporal expression profiles. The underlying hypothesis is that regulatory elements responsible for this shared transcriptional activity are enriched within the promoter sequences of such a set of genes.[1-3] Within the genes, these searches for the regulatory elements are often limited to the conserved non-coding sequences (CNS), which although comprising only 1-2% of the genome are apt to contain sequences involved in transcription regulation.[3-7] Several studies have already

been undertaken to identify motifs involved in tissue-specific gene expression.[8-10] However, many of these studies only involved *in silico* (computer) analysis and provided no experimental validation. The necessity for some sort of 'wet-bench' derived supportive data however becomes apparent from studies like those of Bajic et al., who found that some of the programs in use perform even worse than random guessing.[11]

In our study we first searched for transcription factors and *cis*-regulatory modules responsible for restricting gene expression to the heart's unique muscle cells, the cardiomyocytes. Secondly, we performed a search for transcriptional elements involved in a process that is a unique characteristic for the cardiomyocytes, the response to continuous electric stimulation. For several known cardiac-specific genes it is known that a relative short stretch of 5'-flanking sequence is sufficient to drive cardiac-specific expression. Often, a few hundred base pairs are enough to restrict their expression to the heart.[12-15] However, there are instances known of key cardiac enhancers located further upstream. Restriction of ACTC1 expression to the heart depends for example on an enhancer located 2 kb upstream of the TSS.[15] Taking this into account, we focused our search for regulatory elements at first on the region stretching up to 2 kb directly upstream of the TSS.

Previously we investigated genome-wide the effects of electric stimulation on gene expression and showed that several cardiac-specific genes are activated by electric pulsing independent from contractions.[16] The identification of *cis*-regulatory elements and proteins involved in mediating the response of the cardiac transcriptome to electric signaling will bring valuable information on one of the basic mechanisms contributing to the formation and maintenance of the heart. We hypothesized that a shared regulatory mechanism at the transcription level could be underlying the increase in gene expression of pulsing-responsive genes. Using the aforementioned approach for finding regulatory motifs responsible for a shared transcription program could thus potentially also provide valuable insights in the heart's response to electric signals.

Materials and methods

Data collection and preprocessing

Based on the gene expression databases of SYMATLAS/BioGPS and REFEXA (http://157.82.78.238/refexa/main_search.jsp), we composed 5 sets of human genes, expressed specifically in either, adipocytes, endothelium, heart, skeletal or smooth muscle.[14] An additional sixth set was composed, consisting of 500 random human genes. The composed sets of tissue-specific genes are supplied as supplementary data (supplementary file 1). For the tissue-specific genes and their corresponding mouse orthologues, we downloaded from the ENSEMBL35 database the genomic sequence of the entire gene, plus 10kb up- and down-stream of the gene.[17] The human and mouse sequences were pair-wise aligned with AVID and regions were considered to be conserved when more than 75% was identical within a sliding window of 100bp.[18] Further analysis was restricted to the genomic sequences and conserved regions located within 2kb up- and 0.5kb down-stream of the transcription start site, an approximation of the promoter (supplementary file 2). Moreover, parts that overlapped with adjacent genes were discarded.

Selection of regulatory motifs

Known motifs - For the detection of 'known' motifs we used the positional weight matrices (PWMs) of the TRANSFAC Pro 9.3 collection, thereby only employing the 557 PWMs for vertebrate factors.[19, 20] As background model for motif searches we used the third-order Markov chain background model (BG), supplied by the developers of MOTIFSCANNER[21] This model was computed using the sequences between -2000 bp and +200 bp relative to the TSS of all human genes deposited in ENSEMBL. An intermediate set (Set 1) of TRANSFAC-derived PWMs was assembled by selecting those that had an occurrence within the CNS according to MOTIFSCANNER (prior=0.2, double strand=yes, background=BG).

De novo motifs - Potential *de novo* PWMs/ motifs within the CNS were searched for using the Gibbs sampling tool MOTIFSAMPLER (prior=0.2, double strand=only forward, background=BG, number motifs=1, number runs=100, width= 3-12bp).[21] Per tissue, PWMs were clustered based on width and one representative PWM was

preserved. Clustering and selection was then repeated for all the preserved PWMs for each tissue. For the PWM-clustering we used graph-based clustering with MCL (I=2.2) and Kullback-Leibler, computed with MOTIFCOMPARISON (threshold=5, maximum allowed shift=0) from the INCLUSIVE software package, as measure of divergence.[22, 23] The obtained motifs of this clustering were assembled in a second intermediate set of motifs (Set 2).

Computational identification of tissue-specific motifs

Using MOTIFSCANNER (prior=0.2, double strand=yes, background=BG) we calculated for each PWM in Set1 and Set2 the number of instances in the CNS. Then we analyzed per list of tissue-specific genes whether the frequency was significantly larger than one would expect based on the results for the list of random genes. All PWMs from Set1 and Set2 with a statistically over-represented (SOR)-score ≤ 0.1 were merged into a combined set (Set3). To remove redundant motifs a final list (Set4) was created by discarding de novo PWMs from Set3 that matched with the prior selected TRANSFAC PWMs (Set1). A de novo PWM was considered as matching a TRANSFAC PWM based on 3 criteria: (1) the de novo PWM had an instance that overlapped at least for 75% with an instance of the TRANSFAC PWM, (2) the Kullback-Leibler divergence between the PWMs as computed with MOTIFCOMPARISON (threshold=0.5, maximum allowed shift=2), was less than 0.5 and (3) the logos of the PWMs looked similar after visual inspection.

Computational identification of binding sites and cis-regulatory modules

Binding sites for PWMs in the list of Set4 were located in the complete sequences (irrespective of conservation) using MOTIFSCANNER (prior=0.2, double strand=yes, background=BG). Again the observed number of instances per tissue was compared with the number expected based on the results for the set of random genes. For each tissue only the binding sites of PWMs with a SOR ≤ 0.1 were retained. The predicted binding sites were used to identify cis-regulatory modules. Therefore, we constructed a co-localization network $G=(N,E)$ where N is the set of PWMs and E is a set of weighted edges between them. The weights were computed from the MOTIFSCANNER results per tissue as the probability that the two PWMs

have an instance within 200 bp of each other. Clusters in this network are identified by graph-clustering with MCL (I=1.6). Retained were the heterogeneous clusters that grouped together 2-7 distinct PWMs and for which we could confirm a clustering of the associated binding sites after visual inspection of the MOTIFSCANNER results with TOUCAN[21] Cis-regulatory modules could then be identified as regions in the genomic sequences that were densely populated by instances of the PWMs in one of such PWM clusters.

Computational identification of pulse-responsive regulatory motifs

To identify cis-regulatory elements with a potential role in the transcriptional response of the heart muscle to electric stimulation, we employed the microarray-data from our study in cultured cardiomyocytes.[16] This dataset is accessible under NCBI GEO series accession number GSE15856.[24]

First genes were selected that were up-regulated in cells exposed to electric pulsing. To allow for a larger number of cardiac genes, genes with $p > 0.1$ were selected. Next all genes were filtered out that were down-regulated by addition of blebbistatin. The up-regulation of these genes in the pulsed samples may have been the result from an increase in stretch activity and not from the electric pulses themselves. The reference set was composed of genes of which the expression was not increased by pulsing. From both gene-sets we selected the cardiac-specific genes by comparing the gene-expression levels in human cardiac ventricle with that in a set of 64 other human tissues.[25] By using this dataset we incorporated the assumption that overall tissue-specificity of genes in human and rat is comparable. Tissue specificity is expressed as the entropy value (Q).[26] Genes with an entropy value (Q)<5 were considered heart-specific (supplementary file 3). Genes that occurred in both sets due to results from different probe-sets were omitted from analysis. Similar to the previous analyses overrepresented TRANSFAC motifs within the promoter region -1 kb to +500 bp were identified using MOTIFSCANNER (prior 0.2, both strands). In this analysis rat promoters were investigated. Of the available species for which models are available in MOTIFSCANNER, mouse is the most similar to rat. For this, the third-order model of 1 KB mouse ENSEMBL promoters, provided by the MOTIFSCANNER program was chosen.

Plasmids and constructs

Promoter fragments were obtained through PCR on human genomic DNA. Resulting promoter fragments were cloned into the luciferase reporter vector pGL4.10 (Promega) using standard molecular biology techniques. The promoter fragments used were, *CSRP3*[-458/+77], *MYH6* [-2048/+25] and *HRC* [-965/+73] The human genomic sequences of the selected promoter regions are provided in supplementary file 10. Mutant constructs were generated using the QuickChange II site-directed mutagenesis system (Stratagene, La Jolla, CA USA).

Cell culture and transfection

Cultures of primary neonatal rat ventricular cardiomyocytes were prepared and maintained as described earlier.[27] Transfections were performed using FugeneHD (Roche Applied Biosystems) according to manufacturer's instructions.

Luciferase assays

Luciferase was measured 3 days post transfection using the Dual-Glo reagent (Promega, Madison WI U.S.A.) on a Glomax 96 well luminometer (Promega, Madison WI USA).

Electrophoretic mobility shift assay (EMSA)

EMSA was performed using biotin labeled probes (Eurogentec, Liege Belgium) and a LightShift chemiluminescence EMSA kit (Thermo Scientific, Rockford IL USA). Oligonucleotide sequences of the probes used for EMSA are listed in supplementary file 4.

Results

Predicted cis-regulatory elements and modules

Based on gene-expression data from publicly accessible databases, lists were composed consisting of genes which' expression is restricted to either cardiomyocytes, skeletal muscle, smooth muscle, endothelial cells or adipocytes (supplementary file 1). Smooth muscle and skeletal

muscle were included in the analyses in order to distinguish cardiomyocyte-specific elements from those associated with muscle-tissues in general.

Through a computational procedure (see materials and methods) we derived from human-mouse conserved non-coding sequences (CNS) in the promoter regions of the genes some patterns that characterize the DNA-binding preferences of relevant trans-regulatory elements. The patterns – called positional weight matrices (PWMs) – were selected from an existing database (TRANSFAC) or modeled from scratch. The names of the final selection of TRANSFAC PWMs are available in supplementary file 5. The *de novo* PWMs were predicted by scanning the sequences with MOTIFSAMPLER. We chose for a single-stranded search, since a double-stranded search showed a bias toward palindrome-motifs. Out of the 413 *de novo* PWMs that we identified in the CNS, 123 were retained after evaluation of their statistical overrepresentation. The majority of identified *de novo* PWMs have a width of 6 to 9 bp, enabling the complete binding site to be located inside the major groove of the DNA helix for better accessibility to binding proteins. With the employed filtering procedures, we were able to eliminate some of the similar *de novo* PWMs, but not all of them. Some redundancy remained among the *de novo* PWMs. Patterns that repeatedly can be observed among the cardiac-specific *de novo* PWMs are for instance 'GCAGG', 'GGGNN(N)GGG(G)' and their reverse complement. Matching of *de novo*-PWMs with known (TRANSFAC) PWMs, revealed that 28 'de novo' PWMs (23%) correspond to 8 well-known transcription factors, MAZ, UF, NF-kappa, CHCH, SP, MZF, AP and E2F. The matrices of the final selection of *de novo* PWMs are available in supplementary file 6.

The identified PWMs were used to predict the location of *cis*-regulatory elements in the 'complete' promoter sequences, -2 kb to + 0.5 kb of the TSS, including non-conserved regions. The computational procedure (see materials and methods) yielded a catalogue consisting of 5,278 binding sites for 135 different PWMs (supplementary file 7). We searched in more detail for transcription factors with binding sites that repeatedly appear in close proximity of one another within the promoter set of a certain tissue. Clusters of such transcription factor binding sites are candidate *cis*-regulatory modules. After graph-clustering and filtering (see materials and methods), we found 10 groups of motifs with clustered binding sites in heart, smooth muscle and endothelial tissue (supplementary file 8). Most groups contain a mixture of known and *de novo* PWMs.

Regions 0.5kb proximal to the TSS appear to be enriched both with CNS and cis-regulatory elements (see figure 1). Sequence alignment of the tissue-specific genes with their mouse orthologues revealed that CNS make up for 13 percent of the sequence between - 2kb to +0.5 kb of the TSS and most of these CNS are located within 500 bp of the TSS. The alignment approach selected genomic regions that are slightly enriched with regulatory elements. Depending on the tissue, 15.92 ± 2.42 percent of the binding sites of known transcription factors (PWM set S1, see methods) fall within a predicted CNS according to MOTIFSCANNER. A one-sided Wilcoxon signed rank test confirmed that this percentage is significantly larger than the expected 13 percent (p-value=0.03). Moreover, the enrichment increases with 3.42 ± 0.79 percent when only binding sites on the 5'-3' strand are considered. This increase is also statistically significant according to a one-sided Wilcoxon signed rank test (p-value=0.03).

Comparison of predicted tissue-specific regulatory programs

Comparison of the motif-scanning results of the complete sequences (i.e. including the non-conserved regions) for the 5 tissues, revealed similarities between heart, skeletal- and smooth-muscle. SRF emerges strongly across the various types of muscle. HEN1, which is predicted to play a role in transcription of heart genes, represents an E-box binding protein. Such E-boxes can however be targeted by a wide variety of factors. Analogous to the foregoing, endothelial and adipocyte tissues share fewer factors with the other tissues we investigated here, the muscles. Known motifs that seem to be more typical for the heart are NMYC - associated with N-Myc - and muscle initiator sequence 19 - associated with yet unidentified transcription factor(s). Other known factors with binding sites that are shared between heart and the other tissues are AP2 α , AR, E2A, GLI, HIF1, LBP1, MAZR, MEF2, MEF3, MYOD, MZF1, NF κ -B, RSRFC4, SMAD3, SMAD4, SP1 and TAL1.

The analysis of clustered binding sites revealed in all muscles a grouping of many versions of the PWM of SRF. Skeletal tissue also has a grouping with many versions of the PWM of PPARG. All muscle tissues contain a large cluster that includes SP1/TAL1/MYOD. This combination thus seems to be representative for a general muscle program. Both striated muscles, heart and skeletal, contain a cluster including MEF2 and RSRFC4. Heart tissue has an interesting specific cluster, which includes muscle initiator sequence 19 (a motif that was only included in the cardiac-specific list).

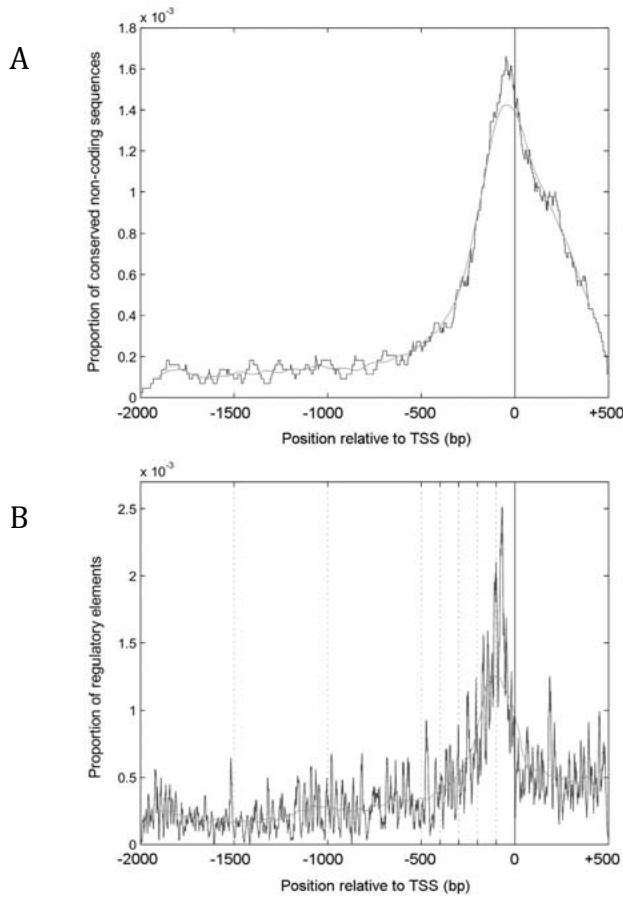


Figure 1. Enrichment of conserved non-coding sequences and transcription factor binding sites in the region surrounding the TSS.

(A) Genomic DNA sequences between -2kb and +0.5 kb of the TSS of human genes and their mouse orthologues were aligned with a sliding window of 100 bp. Sequences not encoding for proteins with >75 percent identity between human and mouse were scored as conserved non-coding sequences (CNS). The number of CNS for each position, normalized over the 2.5 kb region was plotted. The smoothened line depicts the overall trend fitted through all data points. (B) A catalogue was composed of 6,235 computationally predicted binding sites within -2kb and +0.5kb from the TSS of the investigated genes (including non-conserved regions). The graph describes the number of predicted binding sites that cover each position, computed over all sequences and normalized over the 2.5kb region. The trend is indicated by a smoothed line that was fitted through all data points.

Predicted pulse-responsive regulatory program

Based on previously generated micro-array data we selected a set of 12 cardiac-specific genes that were specifically activated by electric stimuli, independent of contraction (supplementary file 2).[28] Using the TRANSFAC matrices, we detected the potential TF-binding sites within the human/rat CNS between 1 kb upstream and 500 bp downstream of the TSS. As a control set, we selected cardiac-specific genes displaying no specific response to electric stimuli. Overrepresentation of known motifs was again determined using a binomial distribution. Motifs identified as being enriched in promoters of pulse-responsive genes are listed in supplementary file 9. The top-ranking motif is E-box 47. Amongst the other motifs identified are several well-known key cardiac transcription factors, GATA, TBX5, MEF2 and MYOG. Especially of the latter a high number of occurrences were found. Interestingly, amongst the motifs identified was also YY1, a factor known to act as a transcriptional inhibitor in muscle.[29] In our analysis we also found LEF1 to be enriched in pulse responsive genes. LEF1 binding sites can be targeted by TCF7L2 the final actor of many pathways leading to the activation of β -catenin that acts in concert with TCF7L2. Due to the limited number of genes available in the set of pulse responsive genes, a prediction of cis-regulatory modules was not attempted.

In vitro verification of identified motifs

In addition to investigating the effects of electric stimuli on the endogenous rat genes, we also performed studies on the effects of electric stimuli on the transcriptional activity of human promoters. For this, neonatal rat cardiomyocytes were transiently transfected with luciferase reporter constructs containing human promoter fragments. Both a -780/+77 promoter fragment of cysteine and glycine rich protein 3 (CSRP3) and a -2028/+30 promoter fragment of alpha myosin heavy chain (MYH6) were found to be strongly activated by electric stimulation (figure 2).

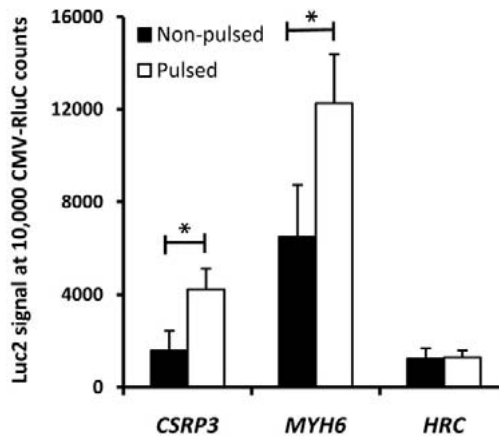


Figure 2. Human promoter fragments of CSR3 and MYH6 are activated by electric stimulation. Primary cultures of neonatal rat ventricular myocytes were transfected with firefly luciferase reporter constructs bearing the human promoter fragments CSR3[-758/+77], MYH6 [-2048/+25] and HRC [-965/+73]. After transfection, cultures were maintained for 3 days with and without electric stimulation (1 ms pulses of 1 Volt/cm at 1 Hz). Signals are normalized to 10,000 counts/s of activity from co-transfected CMV-Renilla luciferase.

Using the CSR3-reporter construct we verified *in vitro* the relevance of the motifs identified in the bioinformatics screening. All the promoter elements that we investigated *in vitro* are conserved across species (figure 3). Using site directed mutagenesis the TFBSs were mutated in separate constructs. Care was given to avoid the introduction of any potential new TFBSs as a result of the mutagenesis. The effect of these mutations is displayed in figure 4. From figure 4 can be seen that transcriptional activity of the CSR3 promoter decreases about 10 fold when disrupting the myogenin site. Mutation of either the TCF7L2 or Sp1 site completely abolished transcriptional activity.

		TCF7L2			
		-465		-458	
Hs	TTTTTCTCTCC	CAAGCT.TCT	TTGAAGAGAT	TATCTAAGTA	
Mm	CACCTCCCCA	CACGCCCTCC	TTGAAGAGAT	TAGTTAAGTG	
Rn	CTTCAGGGTC	TGAACC.TAT	TTGTTGAGGG	TGGCCATTTG	
Cf	TTTCTTTCCC	CAAGTT.TCT	TTGAAGAGAT	TATCTAAGTA	
Bt	CTTATTTTCC	CAAGCT.GCT	TTGAAGAGAT	TATCTAAGTA	

		SP1			
		-285		-276	
HsAGGT	C..AGTGGGG	TGGGGGCCTG	GAGAAATGAT	
Mm	GGC...AGGG	C..ATGGGGG	TGGGGGTGTG	TGGAAGGGAG	
Rn	TTCACCATGC	CGAACTGGGG	TGGAGGTGCA	AAATGTGGAG	
CfAGGT	C..AGGGGAG	TAGGAGCCTG	GAAAAATGAT	
BtAGGT	C..AGGGGAC	TGAGAGCCCG	GAAAAATGAT	

		MYOGENIN			
		-51		-45	
Hs	TTTCAGCCCC	TGATAGCAGT	TGTGTTACTC	ACGTCTCATT	
Mm	TTT.....	...TGGCAAT	TGTGTTACTC	GTGTCTTATT	
Rn	GAGAAA....	TGCTACCAGT	GGA CTC.CAC	GAGGATCAGG	
Cf	TTTCAGCCTC	TGATAGCAGT	TGTGTTACTC	ATGTCTCATT	
Bt	TTTCAGCCTC	TGATAGCAGT	TGTGTTACTC	ATGTCTCATT	

Figure 3. Conservation of TFBSs in CSRP3 promoter across mammalian species. Alignment of CSRP3 5'-flanking promoter sequences for human (Hs), Mouse (Mm), Rat (Rn), Dog (Cf) and cow (Bt). Highlighted are the TFBSs for myogenin, SP1 and TCF7L2. Numeric values indicate distance in base pairs to TSS.

While myogenin and SP1 had both been previously described to play a key role in the transcription regulation of cardiac genes, no similar findings had been done for TCF7L2.[30, 31] To verify whether TCF7L2 indeed binds to the proposed binding site within the CSRP3 promoter we performed an electrophoretic mobility shift assay (EMSA) (figure 5). Biotin labeled probes corresponding to the human genomic sequence containing the TCF7L2 binding site were mixed with nuclear extract from human heart. To verify specificity of binding mutant probes were used, containing the site mutation identical to that of the mutant luciferase construct. In addition a series was included where an

antibody specific for TCF7L2 was added. At the time of performing the experiments the only available antibody was one that binds to TCF7L2 at the region required by the protein to bind to the DNA. As a result, a super-shift could not be detected. However, addition of the antibody resulted in reduced protein binding to the wild type probe, indicating that the binding factor is indeed TCF7L2.

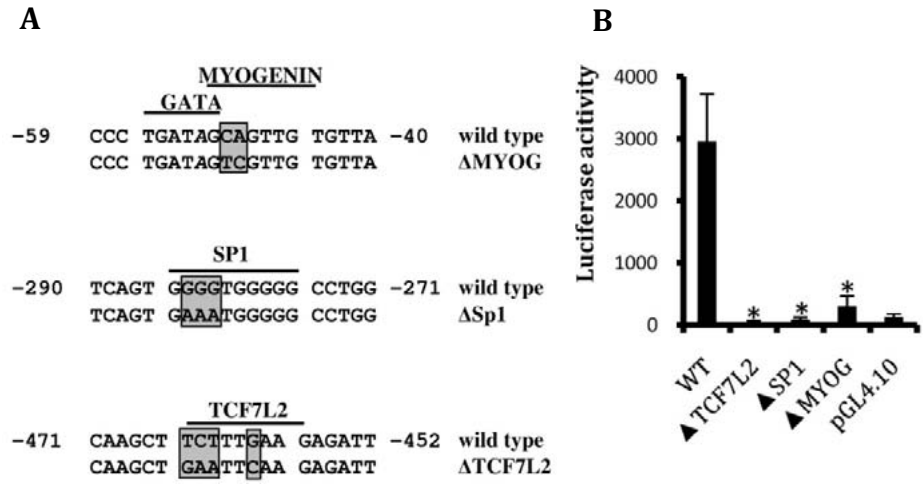


Figure 4. TCF7L2, Sp1 and MYOG binding sites are essential for *CSRP3* transcription.

A: Proposed TFBSs were mutated as indicated in the pGL4.10 firefly luciferase reporter vector containing the wild type *CSRP3*-promoter sequence from -780 to +77 relative to the TSS. The myogenin site partially overlaps with a predicted binding site for GATA factor, the latter was left intact by the mutagenesis.-B: Promoter activity in cardiomyocytes, of pGL4.10 -derived reporter vectors co-transfected with Renilla luciferase (Rluc) vector bearing a CMV promoter. Bars indicate the activity of the wild type (wt) promoter, the three separate mutants and the empty pGL4.10 vector. Signals were normalized to 10,000 Rluc counts/s.

Lane	1	2	3	4	5
Labeled probe	wt	wt	wt	mt	wt
nuclear extract	-	+	+	+	+
unlabeled probe	-	-	++	-	-
Antibody	-	-	-	-	+
α -TCF7L2	-	-	-	-	+

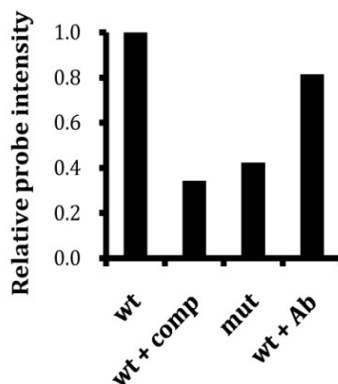
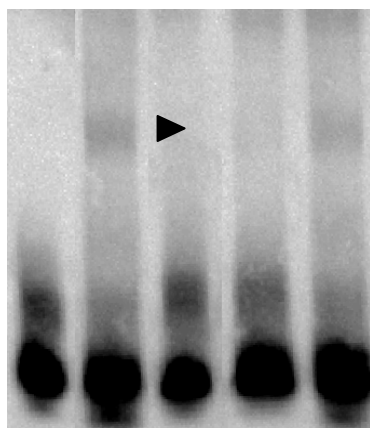


Figure 5. EMSA.

Binding of TCF7L2 to the proposed binding site in the human CSRP3 promoter was investigated with an electrophoretic mobility shift assay (EMSA). Oligonucleotide probes were generated based on the CSRP3 promoter region -476/-447 bp containing the TCF7L2 binding site. Where indicated, unlabeled probe with the wild type sequence was applied in 200 fold the amount of the labeled probe. Graph on the right displays quantification of the indicated band corrected for the background of the individual lanes.

Discussion

In the first part of this study we performed a range of *in silico* analyses to identify motifs and modules involved in orchestrating transcription in a tissue-specific manner. One of the observations we made was that bases in the immediate proximal region of the TSS were the most likely to be part of a conserved non-coding sequence. The observed positional bias is unlikely an artifact of the methodology. Only to a very limited extent, it can be explained by the exclusion of regions

that overlap with adjacent genes: of the 137 2.5kb sequences for the tissue-specific genes, only 15 had an upstream gene that overlapped with the 2.5kb and were henceforth trimmed. The bias was also somewhat to be expected since this is the region closest to the coding sequences which on themselves are in general strongly conserved across species. When screening the sequences, both conserved as well as non-conserved parts, for binding sites, we found an enrichment of regulatory elements within this immediate proximal region. The positional bias of regulatory elements is partially implicated by the use of motifs that were derived from the conserved non-coding sequences.

For a large part, our screening based on matrices for known transcription factors did not reveal much new information on tissue-specific elements. Many of the motifs we identified had already been identified in other studies.[3, 8, 9] The *de novo* matrices that we identified seem to be especially important for regulation in smooth muscle and endothelial tissue, as is indicated by the motif-groups that we found for these tissues which consist of only *de novo* motifs. Since our primary interest was in the heart we did not follow up on investigating the functionality of these proposed motifs and modules.

In addition to identifying regulatory elements involved in tissue specificity we also investigated which *cis*-regulatory elements could be involved in the transcriptional response to electric stimulation.[16] Amongst the motifs identified were E-box47, TBX5, MEF2, GATA1, MYOGENIN, YY1 and LEF1. The function of an E-box 47 site in the MYH6 promoter has been studied in detail by Ojamaa et al.[32] In their study, Ojamaa et al., reported MYH6 to be activated by stretch, based on an absence of contraction due to the addition of 2,3-butanedione 2-monoxime (BDM) or verapamil.[32] However, both verapamil and BDM are known to strongly affect calcium.[33] This together with the observations from our previously reported array study and the fact that MYH6 was found to be not affected by the application of mechanical stretch, lead us to suspect that Ojamaa et al., were in fact looking at a mechanism not resulting from stretch but from an altered calcium signaling such as that induced by electrical pulsing.[34] Many of the motifs identified as being involved in the pulsing response, have also been found to be implicated in hypertrophic responses[35, 36] However, as we reported earlier, we believe that the responses to electrical pulsing and for example phenylephrine induced hypertrophy are quite distinct.[16]

Several findings of the *in silico* analysis are in agreement with literature, but a definite statement about the confidence in the

prediction results should be accompanied by further biological verification, by verification with other tools on the same input data and by verification of our methodology on other input data. The latter will provide an idea about the sensitivity of the methodology to random variations in the inherently incomplete input data, such as the choice of the genes in the random list, choice of the genes in the tissue-specific lists, decision on sequence boundaries, selection of known PWM database (TRANSFAC versus JASPAR) and use of a particular background model. A reliable background model is for instance of uttermost importance.[46-47] The genomic region within 0.5kb from the TSS is rich in G/C content and will naturally be seen as 'awkward' by a background model that is computed from sequences with only normal G/C content.[48] Possibly, the methodology can be made more robust to random variations by controlling the false discovery rate when classifying multiple PWMs as statistically overrepresented, using Benjamini-Hochberg's method.

Next to sensitivity to input data, we should discuss another shortcoming of the *in silico* analysis: the identified regulatory programs are definitely incomplete. Cooperative binding and the role of co-factors were neglected in this study[34] Binding sites for the more ubiquitous expressed SRF emerged from our analyses as being enriched in the promoters genes of 'muscle genes'. For its activity however, SRF often requires the actions of additional factors such as GATA4 and Nkx2-5 in heart and myocardin in heart and smooth muscle.[35-37] GATA4 and especially NKX2-5 are more cardiac-restricted, and expressions of isoforms of myocardin are limited to smooth muscle or cardiomyocytes.[38] Thus although SRF does not directly provide tissue specificity on itself, it can be the key docking site for factors bringing about the eventual tissue specificity. Further limitations of the *in silico* analysis become apparent when one realizes that in cardiomyocytes, the highly cardiac-specific transcription factor Nkx2-5 can also bind to some of the SRF binding sites.[39] This implies that the enrichment of SRF in heart could also be an indication for an increased involvement of NKX2-5.

In the process of verifying *in vitro* the importance of identified motifs in cardiomyocyte transcription we verified for the first time the role of several transcription factors in the transcriptional regulation of the human CSRP3 promoter. While SP1 and myogenin have both been implicated previously to have a direct role in transcriptional regulation of cardiac-specific genes, this has not been the case for TCF7L2. We show here that a mutation of a TCF7L2 binding site within the CSRP3

promoter completely abolished its transcriptional activity in cardiomyocyte cultures. These observations by themselves did not directly confirm a role of TCF7L2 in mediating the transcriptional response to electric stimuli, but even so interesting, they did reveal that TCF7L2 is a key transcription factor for *CSRP3* and possibly more genes with a cardiac-enriched expression profile.

Overrepresentation of the TCF7L2 binding site motif was previously found to be correlated to responsiveness to WNT-signaling.[40, 41] In relation to the heart Wnt/ β -catenin signaling plays a dual role. From experiments in *Xenopus*, Afouda et al. found that Wnt/ β catenin signaling could inhibit cardiogenesis by inactivating GATA4, while GATA4 dependent activation of Wnt11 activated cardiogenesis.[42] Cardiogenesis can be stimulated by Wnt/ β -catenin through Wnt3a and Wnt3 or co-expression of β -catenin and Lef1.[43] In another study, it was shown that exercise training of rat skeletal muscle led to an activation of β -catenin through inactivation of GSK-3 β , via activation of Dvl and Akt.[44]

Taking the aforementioned into consideration, it appears that β -catenin/TCF7L2 signaling, can be activated by fluxes in the electrical potential of cells and still occur in absence of the normally evoked mechanic stretch activity. It should be noted however, that a LEF1/TCF7L2 site was not detected in the promoters of all the pulse-responsive genes. For a complete understanding of the role of LEF1/TCF7L2 and the other pathways involved, a full study would be required to elucidate the mechanisms of the interplay between electric signaling and the mechanic process of contraction.

Although a broad genomics-based analysis as we performed here provides useful insights, it has some serious limitations. Our investigation shows once more that transcriptional regulation is a very intricate process, varying strongly from gene to gene and involving many proteins acting via highly dynamic networks of interactions. In order to fully understand regulatory programs at the transcriptional level, extensive proteomic studies will have to be involved. Large scale studies on co-regulation of genes and protein-protein interaction between transcription-factors are well on the way, but genome wide studies combining proteomics and genomics with resolutions on the individual promoter scale are still yet to come.[45, 46]

Supplementary data

- Supplementary file 1. List of genes used for identifying tissue-specific motifs.
- Supplementary file 2. Promoter sequences-2000bp to +500 bp of genes used in identifying tissue-specific motifs.
- Supplementary file 3 List of genes up-regulated by electric pulsing
- Supplementary file 4. EMSA probe sequences
- Supplementary file 5. List of TRANSFAC PWMs.
- Supplementary file 6. Matrices of identified de novo PWMs.
- Supplementary file 7. Predicted binding sites, tissue-specific motifs
- Supplementary file 8. Candidate cis-regulatory tissue-specific modules.
- Supplementary file 9. Names of pulse responsive motifs.
- Supplementary file 10. Sequences of human promoter fragments used in the luciferase reporter vectors.

References

1. Roeder, R.G., Transcriptional regulation and the role of diverse coactivators in animal cells. *FEBS Letters*, 2005. **579**(4): p. 909-915.
2. Muller, F., M.A. Demeny, and L. Tora, New Problems in RNA Polymerase II Transcription Initiation: Matching the Diversity of Core Promoters with a Variety of Promoter Recognition Factors. *J Biol Chem.*, 2007. **282**(20): p. 14685-14689.
3. Pennacchio, L.A., et al., *Predicting tissue-specific enhancers in the human genome*. *Genome Res.*, 2007: p. gr.5972507.
4. Dermitzakis, E.T., et al., Numerous potentially functional but non-genic conserved sequences on human chromosome 21. *Nature*, 2002. **420**(6915): p. 578-82.
5. Loots, G.G., et al., Identification of a coordinate regulator of interleukins 4, 13, and 5 by cross-species sequence comparisons. *Science*, 2000. **288**(5463): p. 136-40.
6. Wasserman, W.W. and J.W. Fickett, *Identification of regulatory regions which confer muscle-specific gene expression*. *Journal of Molecular Biology*, 1998. **278**(1): p. 167-181.
7. Xie, X., et al., Systematic discovery of regulatory motifs in human promoters and 3[prime] UTRs by comparison of several mammals. *Nature*, 2005. **434**(7031): p. 338-345.
8. Smith, A.D., P. Sumazin, and M.Q. Zhang, Identifying tissue-selective transcription factor binding sites in vertebrate promoters. *PNAS*, 2005. **102**(5): p. 1560-1565.

9. Smith, A.D., P. Sumazin, and M.Q. Zhang, *Tissue-specific regulatory elements in mammalian promoters*. Mol Syst Biol, 2007. **3**.
10. Aerts, S., et al., *Computational detection of cis -regulatory modules*. Bioinformatics, 2003. **19**(90002): p. ii5-14.
11. Bajic, V.B., et al., *Promoter prediction analysis on the whole human genome*. Nat Biotechnol, 2004. **22**(11): p. 1467-73.
12. Durocher, D., et al., The cardiac transcription factors Nkx2-5 and GATA-4 are mutual cofactors. Embo J, 1997. **16**(18): p. 5687-96.
13. Molkentin, J.D., S.M. Jobe, and B.E. Markham, alpha-myosin Heavy Chain Gene Regulation: Delineation and Characterization of the Cardiac Muscle-specific Enhancer and Muscle-specific Promoter. J Mol Cell Cardiol., 1996. **28**(6): p. 1211-1225.
14. Minty, A. and L. Kedes, Upstream regions of the human cardiac actin gene that modulate its transcription in muscle cells: presence of an evolutionarily conserved repeated motif. Mol Cell Biol, 1986. **6**(6): p. 2125-36.
15. Lemonnier, M. and M.E. Buckingham, Characterization of a cardiac-specific enhancer, which directs alpha-cardiac actin gene transcription in the mouse adult heart. J Biol Chem., 2004. **279**(53): p. 55651-8.
16. Martherus, R.S.R.M., et al., Electric signals affect the cardiomyocyte transcriptome independent of contraction. Physiol Genomics., 2010.
17. Hubbard, T.J.P., et al., *Ensembl 2009*. Nucleic Acids Res., 2009. **37**(suppl_1): p. D690-697.
18. Bray, N., I. Dubchak, and L. Pachter, *AVID: A global alignment program*. Genome Res, 2003. **13**(1): p. 97-102.
19. Stormo, G.D., *DNA binding sites: representation and discovery*. Bioinformatics, 2000. **16**(1): p. 16-23.
20. Matys, V., et al., TRANSFAC and its module TRANSCOMP: transcriptional gene regulation in eukaryotes. Nucleic Acids Res, 2006. **34**(Database issue): p. D108-10.
21. Aerts, S., et al., *Toucan: deciphering the cis-regulatory logic of coregulated genes*. Nucleic Acids Res, 2003. **31**(6): p. 1753-64.
22. Coessens, B., et al., INCLUSive: A web portal and service registry for microarray and regulatory sequence analysis. Nucleic Acids Res, 2003. **31**(13): p. 3468-70.
23. Enright, A.J., S. Van Dongen, and C.A. Ouzounis, *An efficient algorithm for large-scale detection of protein families*. Nucleic Acids Res, 2002. **30**(7): p. 1575-84.
24. Barrett, T., et al., NCBI GEO: mining tens of millions of expression profiles--database and tools update. Nucleic Acids Res., 2007. **35**(suppl_1): p. D760-765.
25. Roth, R.B., et al., Gene expression analyses reveal molecular relationships among 20 regions of the human CNS. Neurogenetics, 2006. **7**(2): p. 67-80.
26. Schug, J., et al., Promoter features related to tissue specificity as measured by Shannon entropy. Genome Biol, 2005. **6**(4): p. R33.

27. Martherus, R.S.R.M., V.A. Zeijlemaker, and T.A.Y. Ayoubi, *Electrical stimulation of primary neonatal rat ventricular cardiomyocytes using pacemakers*. Biotechniques, 2010. **48**(1): p. 65-7.
28. Martherus, R.S., et al., Electrical signals affect the cardiomyocyte transcriptome independently of contraction. *Physiol Genomics*, 2010. **42A**(4): p. 283-9.
29. Lee, T.C., et al., Activation of skeletal alpha-actin gene transcription: the cooperative formation of serum response factor-binding complexes over positive cis-acting promoter serum response elements displaces a negative-acting nuclear factor enriched in replicating myoblasts and nonmyogenic cells. *Mol Cell Biol*, 1991. **11**(10): p. 5090-100.
30. Sartorelli, V., K. Webster, and L. Kedes, Muscle-specific expression of the cardiac alpha-actin gene requires MyoD1, CArG-box binding factor, and Sp1. *Genes Dev.*, 1990. **4**(10): p. 1811-1822.
31. Biesiada, E., et al., Myogenic Basic Helix-Loop-Helix Proteins and Sp1 Interact as Components of a Multiprotein Transcriptional Complex Required for Activity of the Human Cardiac alpha -Actin Promoter. *Mol Cell Biol.*, 1999. **19**(4): p. 2577-2584.
32. Munoz, J.P., et al., The transcription factor MEF2C mediates cardiomyocyte hypertrophy induced by IGF-1 signaling. *Biochem Biophys Res Commun*, 2009. **388**(1): p. 155-60.
33. Oka, T., et al., Cardiac-Specific Deletion of Gata4 Reveals Its Requirement for Hypertrophy, Compensation, and Myocyte Viability. *Circ Res*, 2006. **98**(6): p. 837-845.
34. van Oeffelen, L., et al., Detecting cis-regulatory binding sites for cooperatively binding proteins. *Nucleic Acids Res*, 2008. **36**(8): p. e46.
35. Belaguli, N.S., et al., Cardiac tissue enriched factors serum response factor and GATA-4 are mutual coregulators. *Mol Cell Biol.*, 2000. **20**(20): p. 7550-8.
36. Chen, C.Y. and R.J. Schwartz, Recruitment of the tinman homolog Nkx-2.5 by serum response factor activates cardiac alpha-actin gene transcription. *Mol Cell Biol*, 1996. **16**(11): p. 6372-84.
37. Cen, B., A. Selvaraj, and R. Prywes, Myocardin/MKL family of SRF coactivators: key regulators of immediate early and muscle specific gene expression. *J Cell Biochem*, 2004. **93**(1): p. 74-82.
38. Creemers, E.E., et al., Coactivation of MEF2 by the SAP Domain Proteins Myocardin and MASTR. *Molecular Cell*, 2006. **23**(1): p. 83-96.
39. Chen, C.Y., et al., Activation of the cardiac alpha-actin promoter depends upon serum response factor, Tinman homologue, Nkx-2.5, and intact serum response elements. *Dev Genet*, 1996. **19**(2): p. 119-30.
40. Behrens, J., et al., Functional interaction of [beta]-catenin with the transcription factor LEF-1. *Nature*, 1996. **382**(6592): p. 638-642.
41. Hatzis, P., et al., Genome-Wide Pattern of TCF7L2/TCF4 Chromatin Occupancy in Colorectal Cancer Cells. *Mol. Cell. Biol.*, 2008. **28**(8): p. 2732-2744.

42. Afouda, B.A., et al., *GATA transcription factors integrate Wnt signalling during heart development*. Developmental Biology, 2008. **135**(19): p. 3185-3190.
43. Liu, Z., et al., WNT signaling promotes Nkx2.5 expression and early cardiomyogenesis via downregulation of Hdac1. Biochim Biophys Acta, 2009. **1793**(2): p. 300-11.
44. Aschenbach, W.G., et al., Regulation of Dishevelled and beta-catenin in rat skeletal muscle: an alternative exercise-induced GSK-3beta signaling pathway. Am J Physiol Heart Circ Physiol, 2006. **291**(1): p. E152-158.
45. Ravasi, T., et al., An atlas of combinatorial transcriptional regulation in mouse and man. Cell, 2010. **140**(5): p. 744-52.
46. Kawaji, H., et al., Update of the FANTOM web resource: from mammalian transcriptional landscape to its dynamic regulation. Nucleic Acids Res, 2010.

CHAPTER 7

Wnt/Frizzled signaling and electrical stimuli
affect the cardiomyocyte transcriptome via
similar mechanisms

Hilde Laeremans[#]; Ruben S.R.M. Martherus[#]; Torik A.Y. Ayoubi

[#]Equal contribution of both authors

Submitted

Abstract

The Wnt/Frizzled signal transduction pathway is implicated in development and disease. It has already been extensively proven that the Wnt/Frizzled pathway can promote cardiogenesis. Additional preliminary experiments indicate that β -catenin dependent Wnts can promote expansion of cardiac progenitors and differentiation of cardiomyocytes. Based on previous studies we anticipated an involvement of β -catenin signaling in the transcriptional response of cardiomyocytes to electric pulsing. Here, we addressed the role of the Wnt/Frizzled signal transduction cascade in the response to electric signals. We show that the transcriptional response to electric signaling is activated by a β -catenin dependent pathway, which to some extent overlaps with Wnt/Fzd signaling initiated by Wnt5a.

Introduction

The cardiac muscle is one of the three muscle types in the body. Many factors controlling transcription in skeletal and smooth muscles also play a pivotal role in the cardiac muscle. The cardiac muscle, however, is much more difficult to study *in vitro*, because adult cardiomyocytes do not proliferate. In order to approach *in vivo* settings as close as possible, the cardiomyocytes need to align and couple, forming an interlinked electrophysiological syncytium[1] Cultured cardiomyocytes subjected to electric stimulation display increased differentiation associated with a higher degree of organization of sarcomere structures and improved contractile properties[2-5] In the past, many researchers observed that the expression of genes involved in the fetal/hypertrophic program, such as c-FOS, c-JUN, atrial natriuretic factor (NPPA) and brain natriuretic protein (NPPB) were influenced after electrical stimulation.[6, 7] The response of some, however, might be arising from the increased consequential contractions and not the electric stimuli themselves.[8, 9]

Previously, we investigated the genome-wide changes in gene-expression induced by electric stimulation and found that certain cardiac-specific genes were activated solely by electric signals.[3] In a subsequent study, we analyzed the promoters of pulsing-responsive genes to identify regulatory mechanisms underlying the transcriptional

response to electric stimulation. We found that the proximal promoters of genes activated by electric stimuli, independent from stretch, were enriched with target sequences for transcription factors involved in the hypertrophic response such as: MEF2, GATA and TCF/LEF.[10]

TCF/ β -catenin signaling becomes activated in several ways. In response to increased activity and hypertrophy, it is shown to be activated downstream of PKB and Akt-signaling[11]. However, TCF/LEF is also known to be involved in Wnt/Fzd signal transduction cascades, which play an important role in cardiovascular and nervous systems.[12, 13] Interestingly, earlier findings indicated increased expression of several elements of the Wnt-signaling cascade in response to electrical stimulation.[3] The Wnt/Fzd signaling pathway can be divided into the β -catenin dependent pathway and the α -specific or β -catenin-independent pathway. The β -catenin independent pathways consist of a planar cell polarity pathway and can signal through cGMP phosphodiesterases which reduces cGMP, or calcium and a phospholipase-C recruiting NFAT and CamKII.[14] In the β -catenin dependent pathway, binding of Wnt ligands to Frizzled (Fzd) lead to an activation of the protein disheveled which on its' turn inactivates GSK3- β . The decrease in active GSK3- β leads to more β -catenin being translocated to the nucleus and eventually stimulates gene transcription in conjunction with TCF. So far, 19 different Wnts have been identified. They can be subdivided into two classes; the Wnt5a class, which signals via the second messenger calcium and the Wnt1 class, which induces second axis formation and signals via the second messenger β -catenin.[15, 16]

Here, we investigated the role of Wnt/Fzd signaling in the cardiomyocytes' response to electric stimulation. In the model employed, the cells undergo an initial (partial) de-differentiation before (re-)differentiating again to cardiomyocytes displaying the characteristic frequent spontaneous contractions and expressing cardiac-specific proteins.[3] This provides us with an excellent tool to study Wnt/Fzd related interventions and examine their effect on the differentiation of cardiomyocytes and on their contractile activity. Modifying the Wnt/Fzd signal transduction cascade may provide a tool for ameliorating synchronicity and frequency of beating of the cardiomyocytes, which could be beneficial for therapy with stem cells.

Materials and Methods

Cell culture

Primary cultures of neonatal rat ventricular cardiomyocytes were prepared and grown on Permanox chamberslides (Nunc, Langenselbold, Germany) as described earlier.[4] In short; tissue was digested overnight with trypsin at 4°C and the next day with collagenase at 37°C. After preplating, cells were plated at a density of 50,000 cells/cm² in high serum medium supplemented with Ara-C to remove remaining non-myocytes. After another 24 hour incubation in high serum without Ara-C, cells were maintained in low serum differentiation medium, containing 2% horse serum. Where applicable, conditioned media were applied at the time when differentiation medium was applied. The conditioned media were prepared by adding the following supplements to the differentiation medium at the indicated concentrations; verapamil hydrochloride at 1 µM (Sigma-Aldrich, Saint Louis, USA), BayK8644 at 10 µM (Bayer, Mijdrecht, The Netherlands). Wnt3a and Wnt5a conditioned media were prepared as described previously, with the difference that cells were maintained in 2% horse serum.[18, 30, 31] Culture media were refreshed every 24 hours. Where indicated, electric stimulation was performed as described earlier.[32] In short; using pacemakers, cultures were exposed to electric pulses at 1Hz and 2 V/cm bipolar. Electrical stimulation was started at the moment of applying differentiation medium.

Recording of contractile activity

Films were made using a Leica DMRIB microscope equipped with a Hamamatsu digital camera (type C4742-95). The individual images, of which the films were composed, were recorded using a 10X magnification and an exposure time of 0.1 seconds. The final movies displayed the cells in real time. Open Lab version 2.2.5 (Improvision, Waltham, MA, USA) software was used. Using a Thermoplate (Tokai HIT, Shizuoka, Japan), the temperature of the culture medium was maintained at 37 °C.

RT-QPCR

After 3 days in differentiation medium RNA was isolated using the Trizol method (Invitrogen, Merelbeke, Belgium). RNA quantity and quality were analyzed using Nanodrop (Thermo Fischer, Waltham MA USA). First strand cDNA was made using M-MLV reverse transcriptase (Finnzymes, Espoo Finland) starting from 1.0 µg of total RNA. QPCR was performed with qPCR master mix for SYBR Green (Eurogentec, Liège Belgium) in 10 µl reactions in 384 well format on an ABI7900 (Applied Biosystems, Melbourne Australia). Rat peptidylprolyl isomerase-A (*Ppia*) served as the reference gene. Primer sequences for all genes investigated are supplied in the supplementary data table.

Statistical analysis

All values are shown as mean ± S.E.M. Differences between groups were examined for statistical significance using one-way ANOVA with the Bonferroni post-test (Graph Pad Prism). A p-value of less than 0.05 was considered to indicate a statistically significant difference.

Results

Effects of Wnt3a and Wnt5a on cardiomyocyte beating frequency

Electrical pulsing enhances the (re-)differentiation of cultured neonatal cardiomyocytes.[2, 3] To study a the role of Wnt/Fzd signaling in this enhanced differentiation upon pulsing, we cultured neonatal rat ventricular cardiomyocytes (NRVCMs) in various types of media, a standard control medium and media supplemented with either Wnt3a, Wnt5a, the calcium agonist BayK8644, or the calcium antagonist verapamil. Verapamil is used to treat arrhythmias. Irregular occurring depolarisations are counteracted while constant sinus rhythm is still sustained.[17] BayK8644 acts oppositely, it prolongs the time that the L-type calcium channels are opened and with it, the time that the system remains insensitive to new incoming depolarization signals.[17]

The conditioned media containing Wnt3a or Wnt5a, have been tested extensively using our standard testing protocols and proved to exhibit the same results as the recombinant proteins.[18] In addition, series were included in which β-catenin was over-expressed using transient transfection of a β-catenin expression vector.

We observed that the addition of the Wnt conditioned media had a strong effect on the beating frequency of the cardiomyocytes, as seen under a light microscope. Cardiomyocytes maintained in standard medium, contracted with a frequency of around 16 beats per minute (figure 1). This frequency was significantly augmented by addition of Wnt5a, to 44 beats per minute (bpm) and even up to 55 bpm by Wnt3a. In both cases the change in contraction occurred within seconds after adding the Wnt conditioned media.

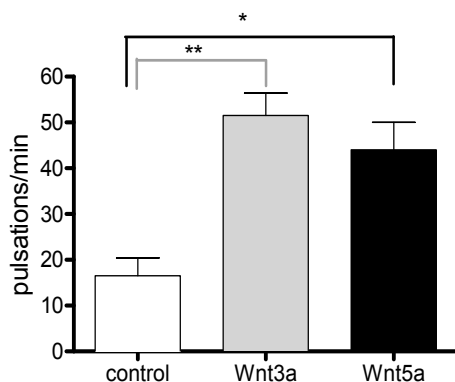


Figure 1. Effect of Wnt proteins on the beating frequency of cultured cardiomyocytes.

Using light microscopy, the contractions of cardiomyocytes were recorded. The cells were maintained medium supplemented with either control, Wnt3a or Wnt5a conditioned medium. The results are represented as a mean of eight independent measurements with SEM (Π $P < 0.05$ and # $P < 0.01$).

Gene expression analysis

For a few selected genes, we investigated whether the use of different media affected their expression at the mRNA-level. The selected genes were those encoding the key cardiac transcription factor GATA4, connexin 43 (*Cx43/Gja1*), which plays a central role in conducting the action potential between cardiomyocytes, alpha myosin heavy chain (*Myh6*) and cysteine and glycine rich protein 3 (*Csrp3*). In our earlier studies we found that transcriptional activity of fragments of the human promoter sequences of *MYH6* and *CSR3* was markedly increased in response to electric stimuli. Furthermore, we showed that a TCF/LEF

binding site in CSRP3 is targeted by TCF7L2 and that this site is crucial for CSRP3 expression in cardiomyocytes.[4, 8, 10]

Using RT-QPCR we found that the expression of the cardiac-specific genes *Myh6* and *Gja1* as well as that of the heart-enriched *Gata4* and *Csrp3* was increased by electric stimulation. Interestingly, supplementation of the medium with Wnt5a resulted in an increased expression of these genes in both the pulsed and non-pulsed cells. The effects of pulsing and Wnt5a appear to be cumulative, the increase in expression of the two combined is comparable to the rise in expression caused by pulsing and Wnt5 individually. Supplementation of the medium with Wnt3a attenuated the expression of all 4 genes in both the pulsed and non-pulsed series. The addition of verapamil led to an increased expression of *Gja1*, *Csrp3*, *Gata4* and *Myh6*. Similar to the observations with Wnt5a, the effects of pulsing and addition of verapamil also seem cumulative. In pulsed cultures containing BayK8644, the expression of all these genes was strongly decreased while in the non-pulsed cells BayK8644 had little to no effect.

Based on our previous research, we anticipated the genes *Csrp3* and *Myh6* to be activated by Wnt/ β -catenin dependent signaling. Indeed, the transcript-levels of both genes were increased as a result of the transient over-expression of β -catenin within the cells. In addition to the aforementioned genes, we also studied the effects of the various conditions on the expression of the hypertrophy markers *Nppa* and *Nppb*. Both genes have been reported to be activated upon pulsation.[6] However, at least for *Nppa*, this activation is likely to be related to the contractile activity and not to the electric stimulation itself.[3, 19] The data of the non-pulsed cells, showed increased expression of *Nppa* and *Nppb* after supplementation with Wnt5a, but decreased *Nppa* and *Nppb* expression after Wnt3a administration (figure 2).

Determination of the Wnt/Fzd signal transduction cascade components

Apart from studying differentiation, we also analyzed the transcription levels of the Wnt receptors. It is known that Wnt3a and 5a are the ligands for Fzd-1 and -2.[20] Figure 3 shows the expression levels for endogenous Fzd-1 and -2. The expression pattern that was observed is similar to that of the differentiation markers (Fig. 2). The relative decrease or increase is more pronounced after pulsation, but the pattern of expression remains the same. Furthermore, Fzd-1 and Fzd-2 also display a similar expression profile.

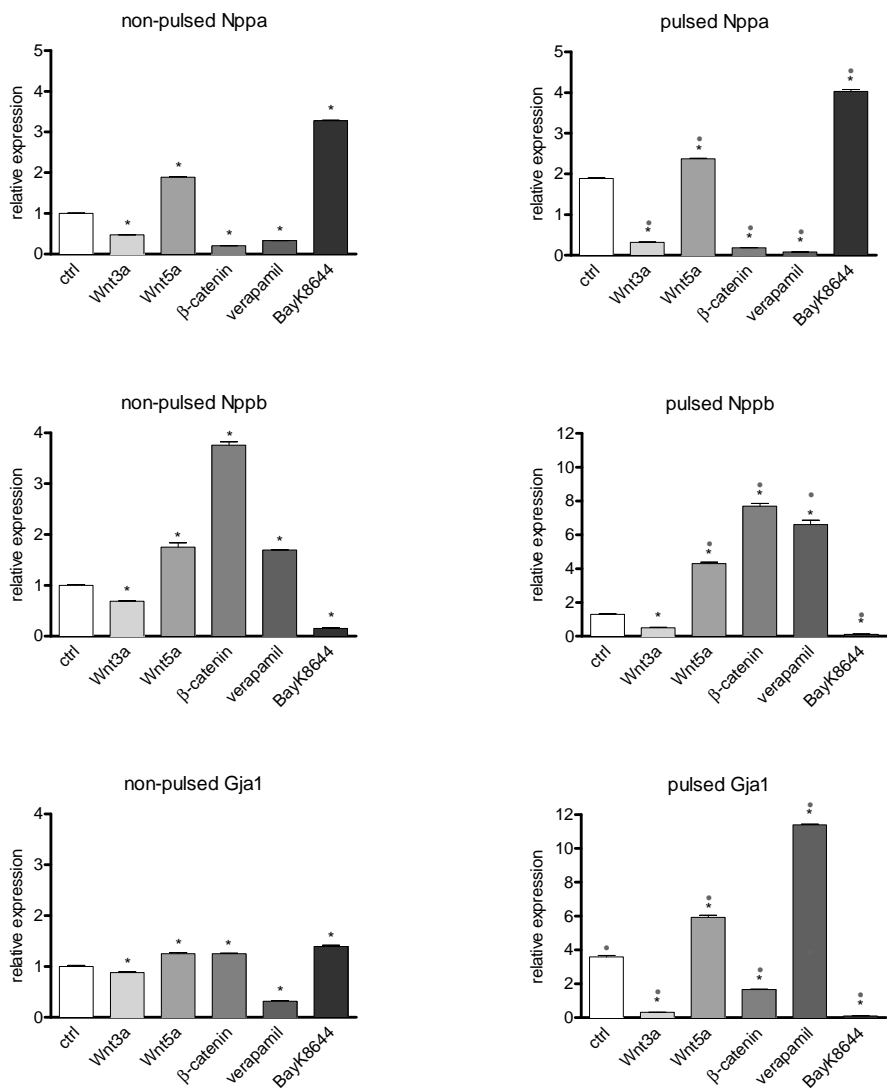


Figure 2: Gene expression analysis (continued on next page)

Left panel are non-pulsed cardiomyocytes and the right panel pulsed. All data are represented as means of six independent measurements and with SEM (* $p < 0.001$ if significantly different from control and • $p < 0.001$ if significantly different from non-pulsed cardiomyocytes).

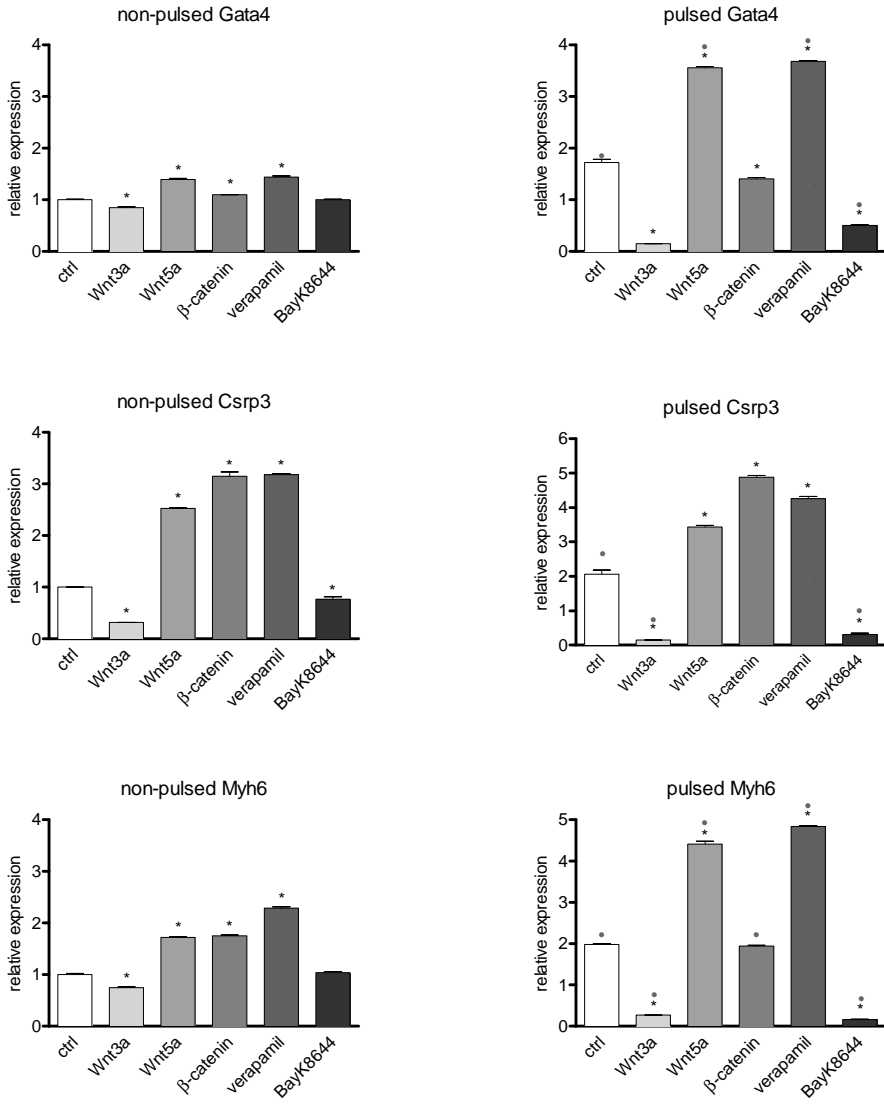


Figure 2: Gene expression analysis (continued)

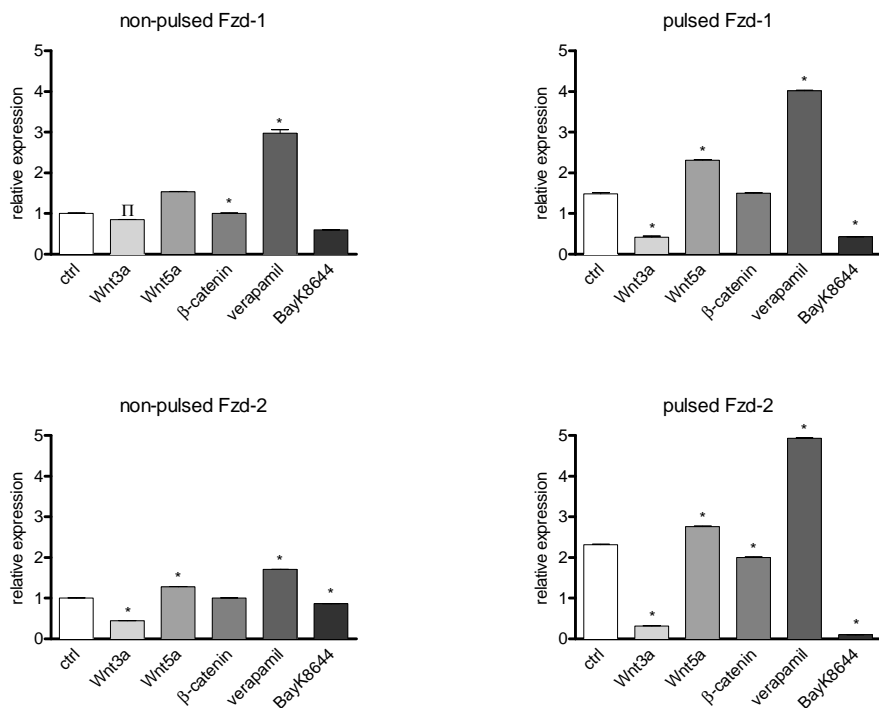


Figure 3: Components of the Wnt/Fzd pathway under different conditions. Left panel are non-pulsed and right panel are pulsed cardiomyocytes. The results indicate means of six independent measurements with SEM. (Π $p < 0.05$ and * $p < 0.001$)

Discussion

The heart consists of different cell types, wherein the fibroblasts, endothelial cells and the cardiomyocytes are the central players. The main functions of the cardiomyocytes are to provide conduction of the action potential and contractile force of the heart. To study the cardiomyocytes *in vitro*, the natural environment of the heart has to be mimicked in the petridish. Cardiomyocytes in culture typically display spontaneous contractions. However, when cardiomyocyte cultures are maintained for longer periods without electric stimulation the contractile activity will be abrogated.[21] Eventually the cells' morphology will change, cells will start to detach and cell death will set in.[22] Therefore, to study cellular mechanisms in the heart, cells have to be pulsed. Previously, we performed in depth studies on the effect of electric stimulation on cardiomyocytes in culture.[3, 10] Similar to others, we found that exposure to electric stimulation enhanced differentiation and improved contractile activity. In this article, we revealed that pulsing activates elements of the Wnt-signaling pathway and the involvement of the TCF/ β -catenin signaling in transmitting the electrical signaling to the transcriptional program.

Here, we found that Wnt5a enhances cardiomyocyte differentiation. Supplementation of culture media with this Wnt activated transcription of genes with a more cardiac-enriched expression profile such as *Cx43/Gja1*, *Csrp3*, *Gata4* and *Myh6*. Moreover, the expression profile of *Myh6*, a known GATA4 target, correlated with that of GATA4.[7]

Wnt3a supplementation, reduced expression of all genes investigated in this study. A similar response for Wnt3a was observed by Liu et al.[23] They found that in later stages of cardiogenesis Wnt3a attenuated differentiation through the inactivation of Nkx2-5 via activation of HDAC1, while in the early stages of cardiogenesis the Wnt3a acted in the opposite way. At the moment when the conditioned media were applied in our experiments, the cells already displayed contractile activity, indicating that the cells had progressed into a more differentiated/late state of development. It is very well possible that if the Wnt3a conditioned media would have been applied in the early stage of (re)-differentiation a different effect could have been observed. An inducing effect of both Wnt3 and Wnt5a on cardiomyocyte contraction has already been described by Toyofuku et al.[24] However, in their experiments, the response to the Wnts depended on the presence of additional fibroblasts. The results were thus affected by cell density, which as we previously demonstrated, even affects

cardiomyocyte differentiation in pure cardiomyocyte cultures.[3] In our experiments, supplementation of the culture media with either Wnt3a or Wnt5a had an immediate effect on the contractile activity of the cardiomyocytes. The frequency of contractions was strongly increased both in electrical pulsed and non-pulsed cells. This response was practically instantaneous and thus had to be initiated by fast acting mechanisms independent from transcription. A plausible explanation of this direct contraction activation could be the Akt mediated activation of L-type calcium channels.[25]

The pattern of gene activation displayed by Wnt5a follows that of the series treated with verapamil. Interestingly, verapamil is a highly potent cardiomyogenesis-inducing agent while at the same time it reduces contractile activity. The down-regulation of *Nppa* in the presence of verapamil is in concordance with the stretch dependent-activation of *Nppa* as proposed by us and others in previous studies.[3, 19]

As expected, based on the presence of the TCF7L2 site within their promoter, we found an up-regulation of *Csrp3* and *Myh6* in the cardiomyocyte-cultures in which β -catenin was over-expressed.[10] Activation of β -catenin suggests an involvement of elements of the β -catenin dependent Wnt/Fzd pathway. Wnt5a induced expression of *Csrp3* and *Myh6* in a way similar to β -catenin, suggesting that this ligand acts on these genes via the β -catenin dependent signaling pathway. Furthermore, our analyses showed that Fzd-2 mRNA-levels were significantly increased by electric stimulation as well as by Wnt5a, indicating a shared feedback mechanism.

Besides the activation of Akt and the β -catenin dependent pathway, Wnt-signaling can act via the β -catenin independent pathway, whereby intracellular calcium levels are increased.[26] Higher calcium levels can also lead to a lower activity of GSK3- β , and thus again to increased β -catenin levels.[27] Earlier, we showed that electric signals in cardiomyocytes activate calcium signaling independent from contraction and the enrichment of TCF7L2 binding sites in the promoters of pulsing responsive genes.[3] We thus believe that this calcium dependent increase in activity of β -catenin is a key switch in the transcriptional response to electric stimuli. Besides allowing for an activation of the β -catenin dependent pathway, the inactivation of GSK3- β also enables GATA4 to become more active. Active GSK3- β results in decreased amounts of nuclear GATA4 by increasing its nuclear export.[28] This would suggest that as a result of electric signaling, the aforementioned increased expression of GATA4 could indeed also result in more GATA4 in the nucleus and thus activation of its target genes.

In short, taken into consideration all the aforementioned observations, the response of cardiomyocytes to electric signaling appears to be transmitted by transcriptional programs which to some extent are shared with the Wnt/Fzd signaling transduction cascade. The β -catenin independent Wnt signaling pathways enable the immediate response. The β -catenin dependent Wnt signaling pathway leads to the activation of for example GATA4, providing a link with the transcriptional machinery to enable long-term adjustments of the cell.

Wnt/Fzd signaling is also known as one of the main transduction pathways in different diseases.[29] Based on the similarities between the physiological responses to Wnt5a and electrical pulsing, we hypothesize that for patients with a disease involving perturbed electric-signaling, the Wnt/Fzd signaling system could well provide an interesting target for therapeutic treatment.

Acknowledgments

This work was supported by BSIK, grant 03033 and the Transnationale Universiteit Limburg.

Supplementary table

Forward and reverse sequences of all primers.

Gene	Forward primer	Reverse primer
<i>Csrp3</i>	CAGTTCCAACAATCCCCAAAG	CAGAGAAGTGGGAAGGGTTGCT
<i>Fzd1</i>	AGCAGCACATTCTGAGGGAGGAG	TCTCTCACCCATCAGTCAGTCCAC
<i>Fzd2</i>	CTGCACTCGTGGAGGAAGTTC	TTCACACGGTGGTCTCTCCAT
<i>Gja1</i>	AAAAAGTTGCTGCTGGACATGA	GAAGGTCGTTGGTCCACGAT
<i>Gata4</i>	CCCTGCCCTTGTTCCAAACAC	TCAGGTATGGGACACATGAGATG
<i>Myh6</i>	ACAGAGTGCTTCGTGCCTGAT	CAGTCACCGTCTTGCCGTTT
<i>Nppa</i>	CATCATGGGCTCCTTCTCCAT	TGTACACAGGATTTGGTCCAATAGT
<i>Nppb</i>	AGGACCAAGGCCTCACAAAA	TTGAGATATGTGTACCTTGAATTT
<i>Ppia</i>	TTCTCTCTTTCACAGAATTATTCCA	CCACCAGTGCCATTATGG

References

1. C. Cannizzaro, N. Tandon, E. Figallo, et al., Practical aspects of cardiac tissue engineering with electrical stimulation. *Methods Mol Med*, 140 (2007) 291-307.
2. E. Holt, P.K. Lunde, O.M. Sejersted, et al., Electrical stimulation of adult rat cardiomyocytes in culture improves contractile properties and is associated with altered calcium handling. *Basic Res Cardiol*, 92 (1997) 289-298.
3. R.S.R.M. Martherus, S. Vanherle, E.D.J. Timmer, et al., Electric signals affect the cardiomyocyte transcriptome independent of contraction. *Physiol Genomics*, (2010).
4. R.S.R.M. Martherus, V.A. Zeijlemaker, and T.A.Y. Ayoubi, Electrical stimulation of primary neonatal rat ventricular cardiomyocytes using pacemakers. *Biotechniques*, 48 (2010) 65-7.
5. Y. Xia, L.M. Buja, R.C. Scarpulla, et al., Electrical stimulation of neonatal cardiomyocytes results in the sequential activation of nuclear genes governing mitochondrial proliferation and differentiation. *Proc Natl Acad Sci U S A*, 94 (1997) 11399-404.
6. P.M. McDonough and C.C. Glembotski, Induction of atrial natriuretic factor and myosin light chain-2 gene expression in cultured ventricular myocytes by electrical stimulation of contraction. *J Biol Chem*, 267 (1992) 11665-8.
7. Y. Xia, J.B. McMillin, A. Lewis, et al., Electrical stimulation of neonatal cardiac myocytes activates the NFAT3 and GATA4 pathways and up-regulates the adenylosuccinate synthetase 1 gene. *J Biol Chem*, 275 (2000) 1855-63.
8. R.S. Martherus, S.J. Vanherle, E.D. Timmer, et al., Electrical signals affect the cardiomyocyte transcriptome independently of contraction. *Physiol Genomics*, 42A (2010) 283-9.
9. D. Frank, C. Kuhn, B. Brors, et al., Gene expression pattern in biomechanically stretched cardiomyocytes: evidence for a stretch-specific gene program. *Hypertension*, 51 (2008) 309-18.
10. R.S.R.M. Martherus, E.D.J. Timmer, S. Vanherle, et al., Identification of promoter elements prerequisite for cardiac specificity and the cardiac response to electric stimuli. Manuscript in preparation, (2010).
11. W.G. Aschenbach, R.C. Ho, K. Sakamoto, et al., Regulation of Dishevelled and beta-catenin in rat skeletal muscle: an alternative exercise-induced GSK-3 β signaling pathway. *Am J Physiol Heart Circ Physiol*, 291 (2006) E152-158.
12. F. Li, Z.Z. Chong, and K. Maiese, Winding through the WNT pathway during cellular development and demise. *Histol Histopathol*, 21 (2006) 103-24.
13. S.W. van den Borne, V.A. van de Schans, A.E. Strzelecka, et al., Mouse strain determines the outcome of wound healing after myocardial infarction. *Cardiovasc Res*, 84 (2009) 273-82.

14. S.P. Shevtsov, S. Haq, and T. Force, Activation of beta-catenin signaling pathways by classical G-protein-coupled receptors: mechanisms and consequences in cycling and non-cycling cells. *Cell Cycle*, 5 (2006) 2295-300.
15. A.J. Chien, W.H. Conrad, and R.T. Moon, A Wnt survival guide: from flies to human disease. *J Invest Dermatol*, 129 (2009) 1614-27.
16. R. Nusse, Wnt signaling in disease and in development. *Cell Res*, 15 (2005) 28-32.
17. D.B. Laurence L. Brunton, Iain Buxton, *Goodman and Gilman's manual of pharmacology and therapeutics*. 2007: McGraw-Hill Professional. 1219.
18. H. Laeremans, S.S. Rensen, H.C.J. Ottenheijm, et al., Wnt/Frizzled signaling modulates the migration and differentiation of immortalized cardiac fibroblasts. *Cardiovasc Res*, (2010).
19. P. Kinnunen, O. Vuolteenaho, P. Uusimaa, et al., Passive mechanical stretch releases atrial natriuretic peptide from rat ventricular myocardium. *Circ Res.*, 70 (1992) 1244-53.
20. H. Laeremans, S.S. Rensen, H.C. Ottenheijm, et al., Wnt/frizzled signalling modulates the migration and differentiation of immortalized cardiac fibroblasts. *Cardiovasc Res*, 87 (2010) 514-23.
21. H.J. Berger, S.K. Prasad, A.J. Davidoff, et al., Continual electric field stimulation preserves contractile function of adult ventricular myocytes in primary culture. *Am J Physiol Heart Circ Physiol* 266 (1994) H341-349.
22. H.J. Berger, S.K. Prasad, A.J. Davidoff, et al., Continual electric field stimulation preserves contractile function of adult ventricular myocytes in primary culture. *Am J Physiol*, 266 (1994) H341-9.
23. Z. Liu, T. Li, Y. Liu, et al., WNT signaling promotes Nkx2.5 expression and early cardiomyogenesis via downregulation of Hdac1. *Biochim Biophys Acta*, 1793 (2009) 300-11.
24. T. Toyofuku, Z. Hong, T. Kuzuya, et al., Wnt/frizzled-2 signaling induces aggregation and adhesion among cardiac myocytes by increased cadherin-beta-catenin complex. *J Cell Biol*, 150 (2000) 225-41.
25. D. Catalucci, D.-H. Zhang, J. DeSantiago, et al., Akt regulates L-type Ca²⁺ channel activity by modulating Cavalpha1 protein stability. *J Cell Biol.*, 184 (2009) 923-933.
26. L.C. Sheldahl, D.C. Slusarski, P. Pandur, et al., Dishevelled activates Ca²⁺ flux, PKC, and CamKII in vertebrate embryos. *J. Cell Biol.*, 161 (2003) 769-777.
27. T. Kanai, T. Nemoto, T. Yanagita, et al., Nav1.7 sodium channel-induced Ca²⁺ influx decreases tau phosphorylation via glycogen synthase kinase-3beta in adrenal chromaffin cells. *Neurochem Int*, 54 (2009) 497-505.
28. C. Morisco, K. Seta, S.E. Hardt, et al., Glycogen Synthase Kinase 3beta Regulates GATA4 in Cardiac Myocytes. *J. Biol. Chem.*, 276 (2001) 28586-28597.

29. V.A. van de Schans, J.F. Smits, and W.M. Blankesteyn, The Wnt/frizzled pathway in cardiovascular development and disease: friend or foe? *Eur J Pharmacol*, 585 (2008) 338-45.
30. K. Willert, J.D. Brown, E. Danenberg, et al., Wnt proteins are lipid-modified and can act as stem cell growth factors. *Nature*, 423 (2003) 448-52.
31. S. Shibamoto, K. Higano, R. Takada, et al., Cytoskeletal reorganization by soluble Wnt-3a protein signalling. *Genes Cells*, 3 (1998) 659-70.
32. H. Fujita, T. Nedachi, and M. Kanzaki, Accelerated de novo sarcomere assembly by electric pulse stimulation in C2C12 myotubes. *Exp Cell Res*, 313 (2007) 1853-65.

CHAPTER 8

FUNCTIONAL ANNOTATION OF HEART-ENRICHED MITOCHONDRIAL GENES GBAS AND CHCHD10 THROUGH GUILT BY ASSOCIATION

Ruben S.R.M. Martherus, Willem Sluiter, Erika D.J.Timmer, Sabina J.V. VanHerle, Hubert J.M. Smeets and Torik A.Y. Ayoubi

Published

Biochemical and Biophysical Research Communications 2010 Nov
12;402(2):203-8

Abstract

Despite the mitochondria ubiquitous nature, many of their components display divergences in their expression profile across different tissues. Using the bioinformatics-approach of guilt by association (GBA) we exploited these variations to predict the function of two so far poorly annotated genes: Coiled-coil-helix-coiled-coil-helix domain containing 10 (*CHCHD10*) and glioblastoma amplified sequence (GBAS). We predicted both genes to be involved in oxidative phosphorylation. Through *in vitro* experiments using gene-knockdown we could indeed confirm this and furthermore we asserted *CHCHD10* to play a role in complex IV activity.

Introduction

Mitochondria are omnipresent organelles involved in various cellular processes such as energy metabolism and regulation of apoptosis.¹ In spite of the ubiquitous nature of mitochondria, their protein composition varies between tissues, largely as a result of differentially expressed nuclear genes.[1] The expression of some components is even restricted to only a single type of tissue. A clear example of this is the cytochrome-c-oxidase (COX) complex/complex IV, the last enzymatic complex in the electron transport chain.[2, 3] *COX6A2* and *COX7A1* both encode isoforms of subunits of complex IV which are only expressed in cardiac muscle and not in any other cell-type.[1, 4, 5]

Large scale proteomics analyses have been performed to dissect the mitochondrial protein composition of several tissues.[6, 7] Apart from registering the presence of known proteins these studies also aimed at identifying 'new' mitochondrial proteins and predicting their function. For these functional predictions proteomics databases have been the principal source, whereas genomics information hardly played any role despite the availability of huge amounts of easily accessible data through depositories such as NCBI's Gene Expression Omnibus.[8] In other studies genomics based data has been widely used to identify functionally related genes and to predict the function of unknown genes.[9, 10] Walker et al., used genomics data from cDNA libraries to predict the involvement of novel genes in prostate cancer. In their approach, called Guilt-by-association (GBA), they scored whether a gene

was transcribed or not in samples from healthy and diseased tissues.[9] Based on the correlation between expression profiles predictions were made on whether genes were functionally linked. The method was thus solely based on co-expression profiles and not on sequence similarities. Since the study of Walker et al. extensive publicly accessible datasets have become available, providing researchers with a rich source of information on the expression of genes across various tissues.[11, 12] We report here on our strategy to use these genomics datasets to predict the function of *CHCHD10/C22ORF16* and *GBAS*, which both have been described as mitochondrial proteins.[13]

Materials and methods

Tissue culture and transfection

HELA cells were grown in 24 well plates in high glucose DMEM (Invitrogen) supplemented with 10% fetal bovine serum. Transfection of siRNA (Dharmacon, Chicago IL. USA) was done using the GeneSilencer system (Genlantis, San Diego USA) according to manufacturer's instructions. Test transfections were performed at 100, 200 and 1000 ng of siRNA per well of a 24-well plate. For all assays (ATP, apoptosis and enzymatic activity) siRNA concentration was set at 1000 ng/ well.

RT-QPCR

RNA was isolated using Trizol reagent (Invitrogen, Madison WI. USA). First strand cDNA was made using M-MLV reverse transcriptase (Finnzymes, Espoo Finland) starting from 1.0 µg of total RNA. QPCR was performed with qPCR master mix for SYBR Green (Eurogentec, Liège Belgium) in 10 µl reactions in 384 well format on an ABI7900 (Applied Biosystems, Melbourne Australia).

Immunohistochemistry

HELA cells were grown on Permanox chamberslides (Lab-Tek, Rochester NY, USA). Cultures were incubated with medium containing Mitotracker® Red CMXRos (Invitrogen, Paisley UK) according to manufacturer's instructions for 20 minutes. Cells were fixated using 4% paraformaldehyde. For staining of CHCHD10 cells were washed 5 times

with PBS-Tween 0.05% and then blocked for 30 minutes with 1% BSA in PBS. Primary antibody, rabbit-anti-human CHCHD10 (Sigma, St. Louis MO, USA) in PBS-Tween 0.05%, was incubated for 45 minutes. After three washes of each 5 minutes with PBS-Tween 0.05%, the secondary antibody, goat-anti-rabbit FITC (Southern Biotech, Birmingham AL USA), was incubated in PBS-Tween 0.05% for 30 minutes. After three washes of each 5 minutes with PBS-Tween 0.05% and two washes with PBS, cells were mounted in 90% glycerol, containing 20 mM Tris-HCl pH 8.0, 0.2% NaN_3 , and 2% 1,4-di-azobicyclo-(2,2,2)-octane (DABCO, Merck, Darmstadt, Germany). Nuclei were counterstained using 0.5 $\mu\text{g}/\text{ml}$ diamidino-2-phenylindole (DAPI; Sigma, St. Louis MO, USA).

ATP assay

Culture medium was replaced with serum free-medium 24 hours after siRNA transfection. After 16 hours of incubation in fresh serum-free medium with applicable supplements oligomycin and/or FCCP were added to a final concentration of respectively 12 μM and 5 μM and the cells incubated for another 10 minutes. Medium was removed and ATP was measured using CellTiter-Glo (Promega, Madison WI. USA) on a Glomax96 luminometer (Promega, Madison WI. USA) according to the manufacturer's instructions.

Apoptosis assay

Apoptosis was screened for using the Caspase-Glo 3/7 assay (Promega, Madison WI. USA). Luciferase measurements were performed with a Glomax96 luminometer (Promega, Madison WI. USA).

Mitochondrial complex measurements

HELA cells were resuspended at a concentration of 10×10^6 per mL in mito buffer consisting of 250 mM sucrose, 10 mM Tris-HCl, 0.2 mM EDTA, pH 7.8, supplemented with one Complete tablet (complete protease inhibitors cocktail; Roche Diagnostics) per 25 mL, snap-frozen in liquid nitrogen and thawed twice to permeabilize the cells. Citrate synthase activity and complex IV assays were performed as previously described.[14] Protein content was determined by the Bio-Rad DC protein assay (Bio-Rad Laboratories) with BSA as a standard.

Results

Identification of heart-enriched mitochondrial genes

A list of 1023 mitochondrial genes was obtained from the dataset of Pagliarini et al.[13] Within this set of genes we searched for genes more restricted to the heart by determining their Shannon entropy values as described by Schug et al.[15] These entropy values are indicative for the enrichment of the mRNA in one given tissue when compared a set of other tissues.¹⁵ The dataset containing the genome-wide expression data across 65 human tissues was obtained from a public accessible dataset via NCBI GEO.[12, 15] A table with all the entropy values of the mitochondrial genes in heart ventricle and atria is provided as an on-line data supplement (supplementary file S1). Amongst the top ranking heart-enriched genes is *CHCHD10/C22ORF16*. In ventricle and atria it ranks at place 6 and 7, respectively.

Besides computer predicted data no information is available on the possible function of this gene. Another poorly annotated gene of which its expression is enriched in heart is glioblastoma amplified sequence (*GBAS*) which ranks 10th in atria and 34th in ventricle. Of both genes we depicted the expression profiles across 65 human tissues in figure 1. As can be seen, the two genes display roughly the same profile with expression enriched in all striated muscles: heart, skeletal muscle and tongue. Overall, the expression levels of *CHCHD10* are much higher than those of *GBAS*.

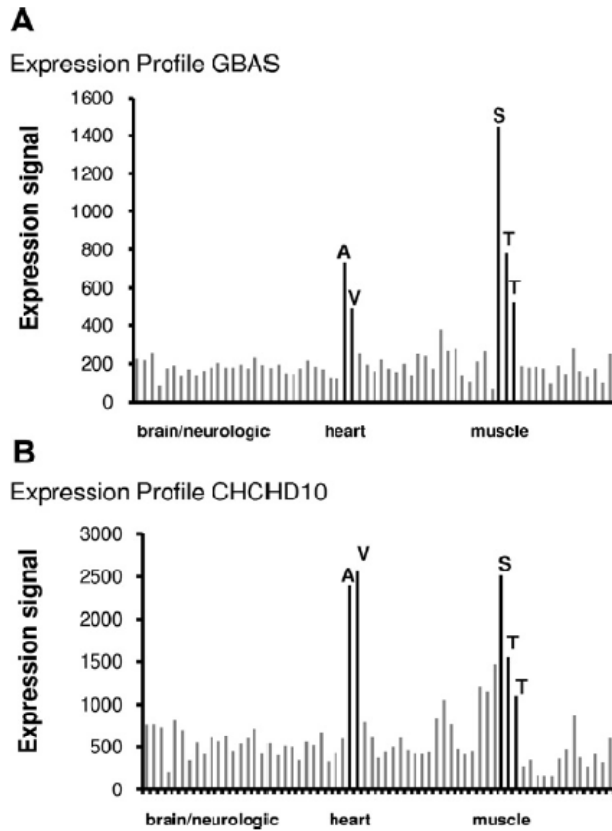


Figure 1: Tissue specificity of mitochondrial genes *GBAS* and *CHCHD10*. Expression profile of *GBAS* (A) and *CHCHD10* (B) across 65 human tissues. Black bars indicate tissues with high expression levels of the genes. V: heart ventricle, A: heart atrium, S: skeletal muscle, T: tongue

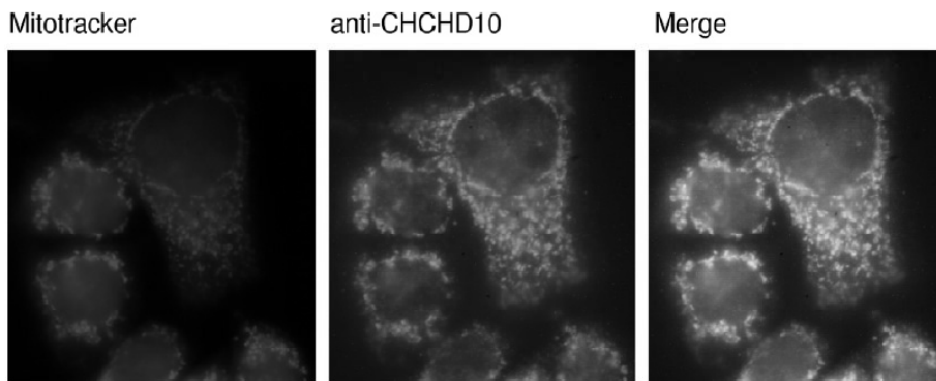


Figure 2: CHCHD10 localizes to mitochondria in HELA-cells.

After culturing and fixation, HELA-cells were stained using with Mitotracker (mitochondrion-selective probe) and an antibody staining specific for human CHCHD10.

Table 1. Top 10 genes with gene expression profile similar to CHCHD10 and GBAS.

Rank	Gene	Pearson corr.	Gene	Pearson corr.
	CHCHD10	1.00	GBAS	1.00
1	UQCR	0.94	CMYA5	0.96
2	NNT	0.94	TTN	0.96
3	DLAT	0.92	ZFP106	0.96
4	ATP5J	0.92	SMPX	0.96
5	NDUFB1	0.92	ACTN2	0.96
6	NDUFB10	0.91	PDLIM5	0.96
7	COX7B	0.90	LMOD2	0.96
8	TMEM143	0.91	PPP1R3A	0.95
9	NDUFS1	0.90	PDE4DIP	0.95
10	MRPL41	0.90	COX6A2	0.95

Genes are ranked by Pearson correlation score calculated from the correlation of all probe sets on the Affymetrix HG U133 plus 2.0 array on 65 human tissues. When represented by more than one probe set, genes were only included once and ranked by the probe set with highest correlation score. Genes in bold font are known mitochondrial genes.

Prediction of GBAS and CHCHD10 function through guilt-by-association

To obtain some indication of gene function of *CHCHD10* and *GBAS* we employed the principle of GBA and compared their expression profiles over 65 human tissues with the profiles of all genes (>38,500) contained within the Affymetrix HGU133 plus 2.0 chip by a simple Pearson correlation test.[12] We calculated the correlation between genes on the basis of their expression data values instead of on the binary presence or absence scores as performed by Walker et al..[9] In table 1 we list the 10 genes with the highest correlation rankings for the two genes separately. We found a high correlation of *GBAS* with several heart-enriched genes with no direct mitochondrial function, whereas amongst the top 10 genes correlating with the expression of *CHCHD10* were only genes with a known mitochondrial function. Amongst them are those encoding: the 6.4kDa-subunit of ubiquinol-cytochrome c reductase (*UQCR*), *COX7B*, subcomplexes of NADH dehydrogenase (ubiquinone) 1 beta (*NDUFB1* and *NDUFB10*) and NADH dehydrogenase (ubiquinone) Fe-S protein (*NDUFS1*).[16-19] Since these genes and *COX6A2* too, encode subunits of oxidative phosphorylation complexes we anticipated both *CHCHD10* and *GBAS*s to be involved in this process. For similar reasons other groups also have been already investigating *GBAS* as possible candidate gene in relation to OXPHOS-disorders.[20]

Protein domains in GBAS and CHCHD10

A screening for protein domains using Interpro based on Refseq mRNA sequences NM_001483.2(*GBAS*) and NM_213720.1 (*CHCHD10*) revealed the presence of a NIPSNAP domain in *GBAS* and a coiled coil-helix-coiled coil-helix (CHCH)-domain in *CHCHD10*. [21,22] The function of all the members of the NIPSNAP super family as well as the NIPSNAP domain itself have remained largely elusive.[22] The CHCH-domain observed in *CHCHD10* has also been found in *COX19* which is involved in the formation of complex IV.[23-25] This and the strong correlation of *CHCHD10*'s expression profile with that of the muscle-enriched complex IV subunit *COX7B*, pointed towards a possible involvement of *CHCHD10* in complex IV function.[26]

Gene knockdown reveals role of GBAS and CHCHD10 in ATP synthesis

To assess a possible involvement of CHCHD10 and GBAS in oxidative phosphorylation we performed gene-knockdown experiments. A simple human model for studying gene expression activity in the heart is not available. However, gene-expression data shows that both CHCHD10 and GBAS are expressed in HELA-cells, thus providing a model of human origin to study the function of the two genes. Using immunohistochemistry we verified the expression of CHCHD10 in HELA-cells and its' localization to the mitochondria within these cells (figure 2).

In our analyses we included Surfeit 1 (*SURF1*) as a positive control since this gene is known to be involved in complex IV assembly.[27] As negative control we employed siRNA specific against firefly luciferase (Luc). Via transfections with siRNA-pools we attained a >50 - 60% gene knockdown for *GBAS*, *CHCHD10* and *SURF1* (figure 3). By culturing in glucose-free medium supplemented with galactose cellular ATP-synthesis is largely restricted to the OXPHOS system. In addition, we supplemented series of cultures with the ATP-synthase inhibitor oligomycin and/or the proton gradient uncoupling agent FCCP to reveal a possible role of the genes in specific steps of the OXPHOS system (figure 4).[28] HELA-cells transfected with siRNA against *CHCHD10* or *SURF1* displayed a significant reduced total ATP-level under all experimental conditions. Also for *GBAS* a significant decrease was found under most conditions, however, in the cultures supplemented with both oligomycin and FCCP the changes in ATP levels in the *GBAS* series just failed to reach significance ($p = 0.054$). We excluded an increase in apoptosis as a cause for the decline in total cellular ATP-levels by analyzing caspase 3 and 7 activity. When compared to Luc siRNA transfected cells we did not observe an increased apoptotic activity due to the siRNA transfections for any of the series (figure 4B).

CHCHD10 is involved in complex IV activity

Based on the observed decrease in ATP formation, the presence of a conserved CHCHD-domain and the strong correlation of its expression profile with that of genes encoding complex IV subunits, we anticipated CHCHD10 to be involved in complex IV activity. Predictions of the function of *GBAS* were less strong. However, recently NIPSNAP1 which like *GBAS* also contains the NIPSNAP-domain, was reported to co-

migrate with intermediates of complex I.[29] We therefore analyzed complex I and IV activity in siRNA transfected cells and in addition also citrate synthase activity was measured to correct for a possible change in mitochondrial copy number (table 2). We observed an increase in citrate synthase activity in the *GBAS* and *CHCHD10* knockdown cells, indicating an increased mitochondrial biogenesis in both series. The change in CS-activity of the SURF1-knockdown series was just short of being significant ($p = 0.055$). After correction for CS-activity it becomes clear that CHCHD10-knockdown indeed leads to a reduced complex IV activity. In cells treated with siRNA for CHCHD10 the remaining complex IV-activity was about half that of the level observed in the *Luc* control series.

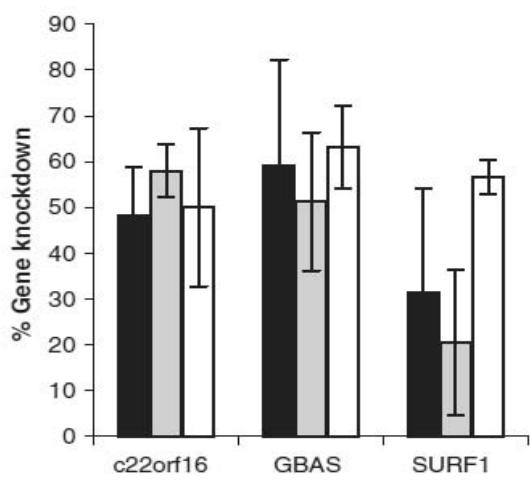


Figure 3: Efficacy of siRNA mediated gene knockdown.

Cultured HELA cells in a 24-well plate were transfected with siRNA at three different concentrations in (100, 200 or 1000 ng per well; respectively: black, grey and white bars). Gene-knockdown was determined through RT-QPCR on samples obtained two days post-transfection; cells transfected with luciferase specific siRNA at identical concentrations served as a reference.

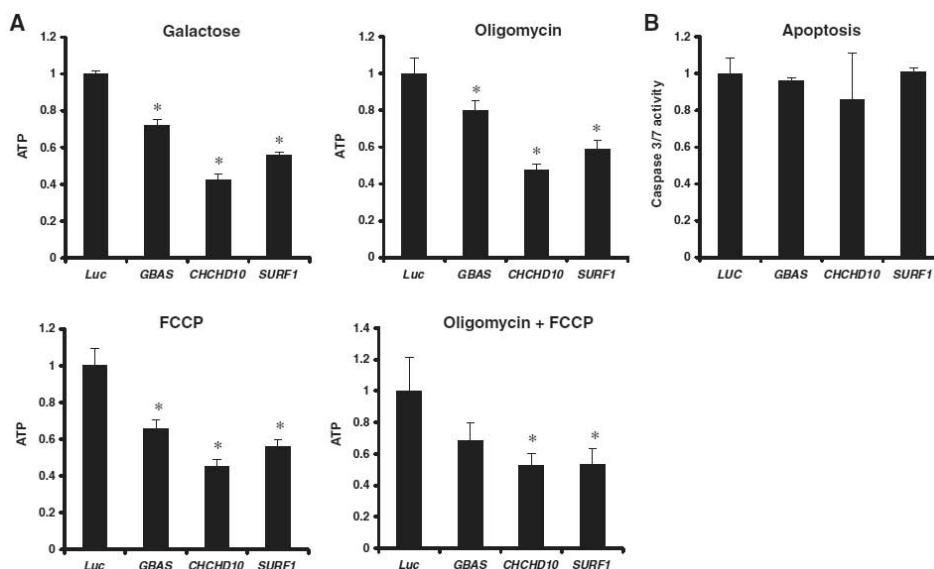


Figure 4. Effect of gene knockdown on cellular ATP content.

HELA-cells were transfected with siRNA specifically for either: *Luc*, *GBAS*, *CHCHD10*, or *SURF1*. One day post transfection cells were incubated overnight in glucose-free medium supplemented with galactose. Additionally several series were treated during the last 10 minutes with oligomycin, FCCP or a combination of these two. (A) For ATP content, total ATP in cell cultures was determined and normalized to the level of citrate synthase activity as a correction for the number of mitochondria. (B) Apoptosis was screened for using a proluminescent caspase 3/7 substrate to verify whether changes in ATP levels were not resulting from increased apoptosis.

Discussion

Various studies have described the protein composition of mitochondria in different tissues.[6, 7] Also *CHCHD10* and *GBAS* have been found through this way to be mitochondrial proteins, however their exact role remained unknown. By correlating gene expression profiles of these genes with those of which the function is known it is possible to predict their function.[9, 10] Where other studies employed a collection of micro-array datasets of various experimental conditions to elucidate gene functions we made use of a dataset containing expression data of genes in a variety of human tissues.[10, 12] Starting from these public accessible datasets we predicted the genes *CHCHD10*

and GBAS to play a role in the process of energy generation by oxidative phosphorylation.[9] Results from our *in vitro* experiments supported this prediction: knock-down of either *CHCHD10* or *GBAS* led to a significant reduction in total cellular ATP levels (figure 4 A). A contributing role of apoptosis to the decreased ATP levels was excluded as none of the siRNAs led to an induction of caspase 3/7 activity (figure 4 B). We found that knock-down of *CHCHD10* led to reduced complex IV activity, indicating that this gene is likely involved in the function of complex IV. Additional experiments however are required to examine how *CHCHD10* exerts its role. In our experiments we also found *GBAS* to be involved in oxidative phosphorylation. Based upon the association of its family-member NIPSNAP1 with intermediates of complex I we investigated a possible role in complex I function.[29] Both complex I and IV activity were not affected by *GBAS* knock-down. Although we describe *GBAS* here as a gene involved in OXPHOS, its specific role in this process remains unclear.

The positive control *SURF1* is involved in complex IV assembly and its knock-down indeed displayed a decreased ATP synthesis, but not the anticipated decrease in complex IV activity.[27] These contradictory data may to some extent result from the limited number of measurements and large variation amongst the data from the complex IV activity measurements. The siRNA knock down level might not have been enough to show the expected effect for *SURF1*, also the protein half-life of *SURF1* could be higher than the other proteins and therefore require a longer time post transfection prior to performing the analysis. A more efficient downregulation could reveal much more clear results, this also holds true for *GBAS* and *CHCHD10*. The use of transgenic or knock-out mice such as has been done in studying *SURF1* could reveal much stronger effects and provide the opportunity to study the role of the proteins in more detail.[30]

Alternatively it is possible that a more stronger effect could be observed for *GBAS* and *CHCHD10* when performing similar experiments in striated muscle cells since the expression *CHCHD10* and *GBAS* is higher within these cells. Skeletal muscle cells are known to contain two distinct groups of mitochondria: intermyofibrillar (IF) and subsarcolemmal (SS).[31] Amongst else, the two mitochondria pools differ in cytochrome-c oxidase activity. While their total cytochrome-c oxidase content is the same, the enzymatic activity is higher in the IF mitochondria.[31] At least for *CHCHD10* it would be interesting to find out whether the protein is perhaps involved in this differential cytochrome-c oxidase activity.

Conclusions

In all we showed here that using the principle of Guilt by Association we successfully identified CHCHD10 and GBAS to play a role in the energy metabolism and in addition identified a functional role of CHCHD10 in complex IV activity. Like CHCHD10 and GBAS many genes are still awaiting their functional anotation. For those uncharacterized genes the bioinformatics approach described here might turn out to be an extremely useful and cost-effective method in providing the primary hint towards the biological process in which they are involved.

References

1. D.T. Johnson, R.A. Harris, S. French, et al., Tissue heterogeneity of the mammalian mitochondrial proteome. *Am J Physiol Cell Physiol*, 292 (2007) C689-697.
2. H.M. McBride, M. Neuspiel, and S. Wasiak, Mitochondria: More Than Just a Powerhouse. *Current Biology*, 16 (2006) R551-R560.
3. P.F. Chinnery and E.A. Schon, Mitochondria. *J Neurol Neurosurg Psychiatry*, 74 (2003) 1188-99.
4. B. Wan and R.W. Moreadith, Structural Characterization and Regulatory Element Analysis of the Heart Isoform of Cytochrome c Oxidase VIa. *J. Biol. Chem.*, 270 (1995) 26433-26440.
5. R.S. Seelan and L.I. Grossman, Structure and organization of the heart isoform gene for bovine cytochrome c oxidase subunit VIIa. *Biochemistry*, 31 (1992) 4696-704.
6. V.K. Mootha, J. Bunkenborg, J.V. Olsen, et al., Integrated Analysis of Protein Composition, Tissue Diversity, and Gene Regulation in Mouse Mitochondria. *Cell*, 115 (2003) 629-640.
7. N. Polyakov, K. Barylyuk, V. Frankevich, et al., Proteomic analysis of heart mitochondria from *Bos taurus*: I. Application of proteomic methods to identification of transmembrane domains of proteins of the internal mitochondrial membrane. *Russian Journal of Bioorganic Chemistry*, 35 (2009) 33-46.
8. T. Barrett, D.B. Troup, S.E. Wilhite, et al., NCBI GEO: mining tens of millions of expression profiles--database and tools update. *Nucleic Acids Res.*, 35 (2007) D760-765.
9. M.G. Walker, W. Volkmut, E. Sprinzak, et al., Prediction of Gene Function by Genome-Scale Expression Analysis: Prostate Cancer-Associated Genes. *Genome Res.*, 9 (1999) 1198-1203.
10. J.D. Wren, A global meta-analysis of microarray expression data to predict unknown gene functions and estimate the literature-data divide. 2009. p. 1694-1701.

11. A.I. Su, T. Wiltshire, S. Batalov, et al., A gene atlas of the mouse and human protein-encoding transcriptomes. *PNAS*, 101 (2004) 6062-6067.
12. R.B. Roth, P. Hevezi, J. Lee, et al., Gene expression analyses reveal molecular relationships among 20 regions of the human CNS. *Neurogenetics*, 7 (2006) 67-80.
13. D.J. Pagliarini, S.E. Calvo, B. Chang, et al., A mitochondrial protein compendium elucidates complex I disease biology. *Cell*, 134 (2008) 112-23.
14. V.P.M. van Empel, A.T. Bertrand, R. van der Nagel, et al., Downregulation of Apoptosis-Inducing Factor in Harlequin Mutant Mice Sensitizes the Myocardium to Oxidative Stress-Related Cell Death and Pressure Overload-Induced Decompensation. *Circ Res.*, 96 (2005) e92-101.
15. J. Schug, W.P. Schuller, C. Kappen, et al., Promoter features related to tissue specificity as measured by Shannon entropy. *Genome Biol*, 6 (2005) R33.
16. M.M. Islam, H. Suzuki, M. Yoneda, et al., Primary structure of the smallest (6.4-kDa) subunit of human and bovine ubiquinol-cytochrome c reductase deduced from cDNA sequences. *Biochem Mol Biol Int*, 41 (1997) 1109-16.
17. C. Ton, D.M. Hwang, A.A. Dempsey, et al., Identification and Primary Structure of Five Human NADH-Ubiquinone Oxidoreductase Subunits. *Biochem Biophys Res Commun.*, 241 (1997) 589-594.
18. T. Emahazion, A. Beskow, U. Gyllenstein, et al., Intron based radiation hybrid mapping of 15 complex I genes of the human electron transport chain. *Cytogenet Cell Genet*, 82 (1998) 115-9.
19. W. Chow, I. Ragan, and B.H. Robinson, Determination of the cDNA sequence for the human mitochondrial 75-kDa Fe-S protein of NADH-coenzyme Q reductase. *Eur J Biochem*, 201 (1991) 547-50.
20. P. Smits, R. Rodenburg, J. Smeitink, et al., Sequence variants in four candidate genes (NIPSNAP1 , GBAS , CHCHD1 and METT11D1) in patients with combined oxidative phosphorylation system deficiencies. *Journal of Inherited Metabolic Disease*, (2009).
21. S. Hunter, R. Apweiler, T.K. Attwood, et al., InterPro: the integrative protein signature database. *Nucleic Acids Res*, 37 (2009) D211-5.
22. E. Seroussi, H.-Q. Pan, D. Kedra, et al., Characterization of the human NIPSNAP1 gene from 22q12: a member of a novel gene family. *Gene*, 212 (1998) 13-20.
23. S. Sacconi, E. Trevisson, F. Pistollato, et al., hCOX18 and hCOX19: Two human genes involved in cytochrome c oxidase assembly. *Biochem Biophys Res Commun.*, 337 (2005) 832-839.
24. M.P. Nobrega, S.C.B. Bandeira, J. Beers, et al., Characterization of COX19, a Widely Distributed Gene Required for Expression of Mitochondrial Cytochrome Oxidase. *J Biol Chem.*, 277 (2002) 40206-40211.
25. K. Rigby, L. Zhang, P.A. Cobine, et al., characterization of the cytochrome c oxidase assembly factor Cox19 of *Saccharomyces cerevisiae*. *J Biol Chem*, 282 (2007) 10233-42.

26. A.B.P. Kuilenburg, J.J. Beeumen, N.M. Meer, et al., Subunits VIIa,b,c of human cytochrome c oxidase. *Eur J Biochem.*, 203 (1992) 193-199.
27. Z. Zhu, J. Yao, T. Johns, et al., SURF1, encoding a factor involved in the biogenesis of cytochrome c oxidase, is mutated in Leigh syndrome. *Nat Genet*, 20 (1998) 337-343.
28. F. Guerrieri, M. Lorusso, A. Pansini, et al., On the mechanism of action of oligomycin and acidic uncouplers on proton translocation and energy transfer in "sonic" submitochondrial particles. *Journal of Bioenergetics and Biomembranes*, 8 (1976) 131-142.
29. H.J.C.T. Wessels, R.O. Vogel, L.v.d. Heuvel, et al., LC-MS/MS as an alternative for SDS-PAGE in blue native analysis of protein complexes. *Proteomics*, 9 (2009) 4221-4228.
30. A. Agostino, F. Invernizzi, C. Tiveron, et al., Constitutive knockout of Surf1 is associated with high embryonic lethality, mitochondrial disease and cytochrome c oxidase deficiency in mice. *Hum Mol Genet.*, 12 (2003) 399-413.
31. A.M. Cogswell, R.J. Stevens, and D.A. Hood, Properties of skeletal muscle mitochondria isolated from subsarcolemmal and intermyofibrillar regions. *Am J Physiol.*, 264 (1993) C383-389.

CHAPTER 9

GENERAL DISCUSSION

A well balanced expression of genes lays at the base of the correct functioning of any organ. A key process enabling this, is transcription. This *pas de deux* of DNA and proteins, controls the spatial and temporal synthesis of the RNA. In this dissertation, we addressed various aspects of transcription in relation to genes with a cardiac-specific expression profile. Where other research typically treated individual genes, we took on the group of cardiac-specific genes as a whole when investigating the mechanisms that give rise to their cardiac specificity. In addition, we investigated how electrical pulses affect the transcriptome of cardiac muscle cells. Finally, we employed transcriptomics data to predict the function of poorly annotated genes with a cardiac-enriched expression profile.

In order to validate our findings from our *in silico* analyses, we generated a number of reporter constructs containing human promoter sequences and verified their activity *in vitro* (**chapter 3**). In other studies on cardiac-specific genes, information is often only provided on which promoter sequences were sufficient for obtaining expression in cardiomyocytes and not whether this expression was specific. To overcome this limitation, we examined the activity of the promoter fragments not only in cardiomyocytes but also in cultures of skeletal muscle and smooth muscle cells. The importance of verifying the activity in related tissues is exemplified by the results we obtained for the CSRP3 promoter. A small fragment was sufficient to obtain high levels of expression in cardiomyocytes, but also displayed activity in both skeletal and smooth muscle cells. *In vivo*, CSRP3 is not expressed in the latter. Extending the promoter fragment to about 5 kb upstream completely abolished activity in the smooth muscle cells. Besides CSRP3, we also provided for the first time information on the human promoters of several other genes, FABP3, MYL3, MYL7, TCAP, and MYBPC1.

A distinctive feature of cardiomyocytes is their continuous contractile activity. Various publications have shown that stretch has a large impact on the transcriptional program in all types of muscle.[1-6] Contractions are the result of fluxes in the intracellular calcium levels, which in turn can be triggered by applying electrical pulses.[7,8] Furthermore, exposing cardiomyocytes *in vitro* to surrogate electrical pulses was shown to affect the expression of cardiac-specific genes like *NPPA* and *MYL2*, and moreover, to improve the overall cardiomyocyte function.[8, 9] This brought us to take on the question, ‘What is the effect of electrical stimuli at the level of the transcriptome? Which genes and associated processes are hereby affected?’

To facilitate studies on the role of electric pulses on the cardiomyocyte transcriptome we developed a culture system in which cells were continuously electrically pulsed using human pacemakers (**chapter 4**). A first analysis revealed that, although the non-pulsed cultures displayed continuous contractions, the application of electric stimuli for several days resulted in a dramatic change in cell morphology and strongly enhanced transcriptional activity of a promoter fragment of human alpha myosin heavy chain (*MYH6*).

Using expression micro-arrays, we further investigated the changes in the cardiomyocyte transcriptome, resulting from the exposure to electric pulsing on a genome-wide scale (**chapter 5**). Similar to others we found that electric pulsing leads to altered cardiomyocyte morphology such as more defined sarcomere structures.[8, 10] At the level of gene expression however, some of our findings did not concur with those described in literature. For example, the expression of *NPPA* (encoding ANF) is often described as a marker to verify that cells are adequately stimulated.[8] Nonetheless, we found that under the conditions employed in our studies, expression of *Nppa* was not increased by electric stimulation. Even more so, we found evidence that *Nppa* is activated by stretch rather than by electrical stimulation, which is in agreement with the finding that ANF levels are increased by passive stretch.[11] These discordant results originate from the fact that often no clear distinction was made between effects derived from the response to electrical stimulation and the response to the evoked stretch. During the course of our studies, Frank et al., reported on the occurrence of a stretch specific program within cardiomyocytes.[3] Our observations lead us to pose the question, 'is there, analogous to the stretch specific program also an electrical-pulsing specific response?'

By applying electrical stimulation in the presence of the contraction-inhibitor blebbistatin, we were able to distinguish the transcriptional responses derived from the electrical signals from those of the stretch normally evoked by these signals. With this approach we have shown that electrical signals indeed activate a transcriptional program 'independent' from stretch and that genes encoding various elements of the calcium signaling were still activated by electrical stimuli in the absence of stretch. This indicates that calcium fluxes play a role in the pulse-response. In addition, we found that the activation of some genes, including *Myh6*, was even stronger when contraction was inhibited. This can be explained by the fact that cyclic stretch has an inhibitory effect on calcium handling.[1]

Furthermore, Myh7 was found to be more stretch dependent. The balance of the myosin heavy chain composition in the heart is of vital importance.[12, 13] This example of the myosin heavy chains shows how at any given time, the complete cardiac transcriptome requires to balance the interplay between the transcriptional programs activated by stretch and electrical signaling. Further research on the interplay between the two programs is necessary to obtain more insight into how a healthy status quo in the heart is maintained.

It is also interesting to see that various genes involved in the fatty acid metabolism are activated by pulsing. Under normal healthy conditions, fatty acids are the primary energy source for the heart while under diseased conditions the heart has to rely on alternative sources such as glucose.[14] This provides an explanation how perturbations in the cardiac electrical signaling can directly affect the metabolism of the heart.

In **chapter 6**, we investigated the promoters of genes specifically expressed in either cardiac, skeletal or smooth muscle. Our initial analysis was aimed at identifying motifs and modules responsible for the tissue specificity. The questions to be answered here were, ‘to what level does the cardiac gene expression program overlap with the programs of the other muscle tissues, and to what extent is it unique to the heart? Is there a common transcription regulatory program that lays at the basis of restricting the expression of genes to the heart?’

Our search for motifs responsible for the tissue specificity of genes, did not result in the identification of many more ‘new’ motifs than those described in the literature that became available during the course of our studies.[15-17] However, it did provide valuable information about the basic architecture of the cardiac-specific genes. The elements most likely responsible for the tissue-specific expression appeared to be enriched in the regions immediate 5’ proximal to the genes. This is in agreement with previous findings that a relatively short promoter fragment, of a few hundred base pairs, was sufficient for driving cardiac-specific transcription of for example *NPPA* and *MYH6*. [18, 19] Furthermore, it supports the idea that the promoter fragments described in chapter 3, which with the exception of TCAP, all encompassed over 0.5 kb of the 5’ flanking region, indeed contain most of the regulatory elements required to obtain the same tissue-specific expression profile as the endogenous gene.

In chapter 1 it was already indicated that co-factors, such as myocardin, are of great importance for distinguishing the

cardiomyocyte transcriptome from that of other types of cells including that of other types of muscle. The bioinformatics-approach used in our search for tissue-specific motifs, did not identify such co-factors. Furthermore, the employed approach failed to yield motifs like Nkx2-5 and GATA4, even though these are known key players in regulating the cardiac transcriptome.[20-22] SRF however, was found as being overrepresented in the different muscle tissues. Both Nkx2-5 and GATA4 are known to be recruited to cardiac promoters as co-factors, acting via SRF.[23-25] Although the identification of statistically overrepresented motifs can provide valuable information, it appears that it is not possible (yet) to completely elucidate the transcriptome regulatory mechanisms in any type of cell by bioinformatics analysis of genomics data alone.

Where Frank et al., identified a stretch specific program, we anticipated that certain genes' transcription is activated by electric pulsing itself.[3] Besides the 'tissue-specific elements', we therefore also screened for regulatory motifs possibly involved in the cardiomyocytes' response to electric stimuli. For this we selected, based upon our microarray data, genes that are specifically activated by electric pulsing. E-box 47 ranked at the top of the list with candidate motifs to be involved in the pulsing-response. Ojamaa et al., reported that this element plays a role in mediating the transcriptional response of MYH6 to stretch.[26] However, the compounds employed in the latter study strongly affect calcium handling. This, together with the data from our studies presented here and findings from others, show that MYH6 is not stretch-responsive and that the 'stretch-response' proposed by Ojamaa et al., is in fact a transcriptional response to altered calcium signaling, as observed in reaction to electrical pulsing.[3, 26, 27] This renders E-box 47 to be a pulsing responsive element as opposed to one involved in modulating the stretch signaling.

Amongst the other motifs we found to be overrepresented in the proximal promoters of pulsing responsive genes was TCF/LEF. Through *in vitro* studies we showed that mutating the TCF7L2 binding site in the human *CSRP3* promoter completely abolished transcriptional activity. Besides being a possible candidate for transmitting the pulse-response to the transcriptome, TCF7L2 thus appeared to play a crucial function in the transcriptional regulation of some genes expressed in the heart.

TCF/LEF is well known as binding partner of β -catenin, a chief player in linking various signaling pathways to the transcriptional

machinery.[28] In relation to the heart, TCF/LEF has been described as being involved in activating *NPPA* in response to phenylephrine induced hypertrophy.[29] In our array-study, we did not find an increased activation of *NPPA* to electric stimuli. But we did observe an up-regulation of several elements of the Wnt-signaling. Wnt signaling has been shown to enhance expression of cardiac genes like connexin-43 via β -catenin.[30] This combined with the foregoing, lead us to investigate whether some relation exists between the gene expression programs activated by electrical stimuli and that activated by Wnt signaling. In **chapter 7** we described how we found Wnt5a and electrical stimulation to affect the transcription in a similar fashion. Evidence for a central role of calcium signaling in the pulse-response of several cardiac genes was obtained through experiments employing the calcium agonist BayK8644 and the calcium antagonist verapamil. Addition of verapamil made the cardiomyocytes more susceptible to the electrical signals applied via the pacemakers. We found that promoter fragments of both *CSRP3* and *MYH6* were activated by pulsing. The promoter of each gene contains a TCF/LEF site, thus potentially enabling recruiting of β -catenin to the transcription initiating complex. As anticipated, both were also found to become activated by over expression of β -catenin.

Our results further suggested a role for β -catenin dependent signaling in the response of these genes to Wnt5a. Interestingly, our analyses also revealed that Wnt5a supplementation as well as electrical pulsing leads to a significant increase in Fzd-2 expression, indicating a shared feedback mechanism.

Addition of either Wnt5a or Wnt3a to the media leads to an instant two fold increase in the frequency of contraction of the cells. This speed at which this response arose indicates that transcription did not play an immediate role in this. The rapid contractions that occurred, again suggest that calcium handling activity was activated. Some of the effects, such as the increase in contractile activity had been reported earlier.[31] However, where the previous study suggested that fibroblasts were required, we show that the effects also occur in pure cardiomyocyte cultures. Based on our findings we concluded that at the transcriptional response to electric signaling is partially mediated through a calcium dependent activation and through activation of β -catenin dependent pathways. Our data strongly suggest that the signaling mechanisms in response to electric pulses and Wnts are linked. Beta-catenin has been known to be of affect muscle development through interaction with MyoD.[32] For this activation TCF/LEF appeared to be not required.[32] We found however that disruption of a

TCF7L2 site resulted in the complete ablation of *CSRP3* transcription. Therefore we concluded that, although β -catenin is likely to play an important role in the response, TCF7L2 is still required. However, a better understanding of their exact interplay in relation to the target promoters will require additional investigations.

Besides utilizing gene expression data to understand cellular responses to various stimuli, gene expression data can also be employed to investigate, or even to some degree predict, the function of individual genes.[33] In **chapter 8** we described how we employed already publicly available gene-expression data to predict the function of two genes, CHCHD10 and GBAS, which lacked any functional annotation short of being found to have a relation to the mitochondria.[34] Micro array data alone were sufficient to present us with the strong suggestion that the genes have a function related to oxidative phosphorylation. Additional 'simple' database searches and the presence of a specific protein domain brought us to suspect CHCHD10 to have a role in complex IV activity. In subsequent *in vitro* experiments we were indeed able to confirm that both GBAS and CHCHD10 have a function in oxidative phosphorylation. For CHCHD10 we were even able to show that the protein indeed has a role related to complex IV. These results demonstrated that bioinformatics analyses provide researchers with a fast and inexpensive way to perform an initial study, based on the results of which additional experiments can be designed. The number of targets to investigate can be greatly decreased in this way, saving costs as well as time. In this era of science, the most important piece of equipment in a lab might thus well be a computer with access to the internet.

Future perspectives

Our findings show once more the importance of *in vitro* model systems to be as true as possible representations of the *in vivo* situation. *In vivo*, all muscle cells are regularly subjected to some form of stretch. For example smooth muscle cells in the wall of the larger arteries receive stretch signals as a result from the variations in blood pressure in the lumen. Skeletal muscles are contracting and stretching as a result of movement. The heart is continuously active, pumping the blood through the body with every beat. Stretch has a considerable impact on the transcriptome and eventual function of the cells.[3, 4, 35] It would therefore be highly desirable to incorporate some form of

stretch stimulation in any study involving an *in vitro* system based on muscle cell cultures.

Also in transfection studies such as we performed, stretch could well affect the outcome.[36] Therefore, based on the aforementioned, a replication of the transfection studies, in both stretch stimulated skeletal and smooth muscle would be advisable. In addition it should be noted that, to be able to make reliable comparisons between various studies of gene-expression in cardiomyocytes or in any other type of muscle for that matter, it is necessary to include detailed statements on contractile status of the muscle cells or tissues employed in these studies. Parameters such as contractile force or shortening fraction and contraction frequency could be highly informative.

Although electrical stimuli have the power to evoke contractions, a combined application of both electrical and mechanical stretch stimulation would provide the best model for studies involving cardiomyocytes. In our experiments we used pure cultures of cardiomyocytes, grown at a relatively low cell-density. This hampers the translation of our results to the *in vivo* situation. Also, interactions between the various types of cells that constitute the heart may affect the outcome. Further studies in to the role of electric signaling in the heart will therefore require the use of models that incorporate these factors. For example, 3D culture systems with multiple cell layers and co-cultures of cardiomyocytes and myofibroblasts or full mixed cell-populations could prove helpful.[37]

For a more comprehensive understanding of the (cardiac)muscle transcriptome, functional genomics data are required. A rich source of information on protein-DNA interactions is ChIP on chip, a technique through which chromatin regions can be identified that are targeted by a specific transcription factor.[38] Such array techniques can also be employed to determine binding specificity of individual transcription factors.[39] Binding specificity can also be investigated on a large scale by using massively parallel sequencing.[40] These approaches resolve binding specificity only on the level of individual factors, additional proteomics methods are still required to identify associating factors and co-factors. But for now no other (technological) solutions have become available to investigate this next level of interactions between the proteome and genome on a large scale.

Further studies on the effects of electrical signaling on the cardiomyocyte transcriptome are clearly needed. Even when the

directly involved signaling mechanisms are not the focus of a study, researchers should be aware of the importance of the electrical component of the heart. After all we showed that it directly affects the expression of genes involved in many key cellular structures and processes: development, the contractile apparatus, intercellular signaling and metabolism.

Pacemaker therapy is widely used to treat patients who suffer from heart-failure. Placement of a pacemaker requires surgical intervention, and considerable time often passes between the moment of diagnosis and the moment at which the device is placed. In the meantime, patients regularly receive treatment in the form of drugs such as diuretics, vasodilators, and inotropic agents, in order to reduce the workload of the heart.[41] However, no direct pharmacological activation of the pulsing-dependent pathways is undertaken at this point. We recommend to pursue in to further studies on manipulating the pulsing-dependent transcription program via pharmacological compounds. Drugs that activate the pathways excited by pulsing might well contribute to an improved outcome for the patients involved.

Our study of the mitochondrial genes exemplifies the functional diversity of mitochondria within a single cell. As posed in chapter 8, we believe that highly enriched expression of both GBAS and CHCHD10 in striated muscle tissue, may well be related to the differences between the inter-filamentary and sub-sarcolemal mitochondria.[42] It will be interesting to find out what the underlying mechanisms are involved in this. The transcriptional activation of these muscle-enriched mitochondrial genes is likely to involve a transcription factor that more restricted to the striated muscle. Proteins like CSR3 which are muscle-specific and present functionalities in both the structure of sarcomeres as well as in transcriptional control could be interesting candidates to study within this context.[43,44]

References

1. B.M. Cadre, M. Qi, D.M. Eble, et al., Cyclic Stretch Down-regulates Calcium Transporter Gene Expression in Neonatal Rat Ventricular Myocytes. *Journal of Molecular and Cellular Cardiology*, 30 (1998) 2247-2259.
2. L.C.G. Campos, A.A. Miyakawa, V.G. Barauna, et al., Induction of CRP3/MLP expression during vein arterialization is dependent on stretch rather than shear stress. *Cardiovasc Res.*, 83 (2009) 140-147.
3. D. Frank, C. Kuhn, B. Brors, et al., Gene expression pattern in biomechanically stretched cardiomyocytes: evidence for a stretch-specific gene program. *Hypertension*, 51 (2008) 309-18.
4. A. Kumar and A.M. Boriak, Mechanical stress activates the nuclear factor-kappaB pathway in skeletal muscle fibers: a possible role in Duchenne muscular dystrophy. *Faseb J*, 17 (2003) 386-96.
5. F. Raymond, S. Metairon, M. Kussmann, et al., Comparative gene expression profiling between human cultured myotubes and skeletal muscle tissue. *BMC Genomics*, 11 (2010) 125.
6. D. Kelly, L. Mackenzie, P. Hunter, et al., Gene expression of stretch-activated channels and mechanoelectric feedback in the heart. *Clin Exp Pharmacol Physiol*, 33 (2006) 642-8.
7. H. Cheng, W.J. Lederer, and M.B. Cannell, Calcium sparks: elementary events underlying excitation-contraction coupling in heart muscle. *Science*, 262 (1993) 740-4.
8. E. Holt, P.K. Lunde, O.M. Sejersted, et al., Electrical stimulation of adult rat cardiomyocytes in culture improves contractile properties and is associated with altered calcium handling. *Basic Res Cardiol.*, 92 (1997) 289-298.
9. P.M. McDonough and C.C. Glembotski, Induction of atrial natriuretic factor and myosin light chain-2 gene expression in cultured ventricular myocytes by electrical stimulation of contraction. *J Biol Chem.*, 267 (1992) 11665-11668.
10. M. Radisic, H. Park, S. Gerecht, et al., Biomimetic approach to cardiac tissue engineering. *Philos Trans R Soc Lond B Biol Sci.*, 362 (2007) 1357-1368.
11. P. Kinnunen, O. Vuolteenaho, P. Uusimaa, et al., Passive mechanical stretch releases atrial natriuretic peptide from rat ventricular myocardium. *Circ Res.*, 70 (1992) 1244-53.
12. J. James, L. Martin, M. Krenz, et al., Forced Expression of alpha-Myosin Heavy Chain in the Rabbit Ventricle Results in Cardioprotection Under Cardiomyopathic Conditions. *Circulation*, 111 (2005) 2339-2346.
13. J.E. Stelzer, S.L. Brickson, M.R. Locher, et al., Role of myosin heavy chain composition in the stretch activation response of rat myocardium. 2007. p. 161-173.
14. B. Wittels and J.F. Spann, Jr., Defective lipid metabolism in the failing heart. *J Clin Invest*, 47 (1968) 1787-94.

15. L.A. Pennacchio, G.G. Loots, M.A. Nobrega, et al., Predicting tissue-specific enhancers in the human genome. *Genome Res.*, (2007) gr.5972507.
16. A.D. Smith, P. Sumazin, and M.Q. Zhang, Identifying tissue-selective transcription factor binding sites in vertebrate promoters. *PNAS*, 102 (2005) 1560-1565.
17. A.D. Smith, P. Sumazin, and M.Q. Zhang, Tissue-specific regulatory elements in mammalian promoters. *Mol Syst Biol*, 3 (2007).
18. D. Durocher, F. Charron, R. Warren, et al., The cardiac transcription factors Nkx2-5 and GATA-4 are mutual cofactors. *Embo J*, 16 (1997) 5687-96.
19. J.D. Molkentin, S.M. Jobe, and B.E. Markham, alpha-myosin Heavy Chain Gene Regulation: Delineation and Characterization of the Cardiac Muscle-specific Enhancer and Muscle-specific Promoter. *J Mol Cell Cardiol.*, 28 (1996) 1211-1225.
20. R. Bodmer, The gene tinman is required for specification of the heart and visceral muscles in *Drosophila*. *Development*, 118 (1993) 719-729.
21. J.L. Sepulveda, S. Vlahopoulos, D. Iyer, et al., Combinatorial Expression of GATA4, Nkx2-5, and Serum Response Factor Directs Early Cardiac Gene Activity. *J Biol Chem.*, 277 (2002) 25775-25782.
22. S. Morin, F. Charron, L. Robitaille, et al., GATA-dependent recruitment of MEF2 proteins to target promoters. *EMBO J.*, 19 (2000) 2046-2055.
23. B. Chen and R.W. Lim, Physical and Functional Interactions between the Transcriptional Inhibitors Id3 and ITF-2b. Evidence toward a novel Mechanism regulating Muscle-specific Gene Expression. *J Biol Chem.*, 272 (1997) 2459-2463.
24. N.S. Belaguli, J.L. Sepulveda, V. Nigam, et al., Cardiac tissue enriched factors serum response factor and GATA-4 are mutual coregulators. *Mol Cell Biol.*, 20 (2000) 7550-8.
25. C.Y. Chen, J. Croissant, M. Majesky, et al., Activation of the cardiac alpha-actin promoter depends upon serum response factor, Tinman homologue, Nkx-2.5, and intact serum response elements. *Dev Genet*, 19 (1996) 119-30.
26. K. Ojamaa, A.M. Samarel, and I. Klein, Identification of a Contractile-responsive Element in the Cardiac alpha-Myosin Heavy Chain Gene. *J Biol Chem.*, 270 (1995) 31276-31281.
27. R.S. Martherus, S.J. Vanherle, E.D. Timmer, et al., Electrical signals affect the cardiomyocyte transcriptome independently of contraction. *Physiol Genomics*, 42A (2010) 283-9.
28. J. Behrens, J.P. von Kries, M. Kuhl, et al., Functional interaction of [beta]-catenin with the transcription factor LEF-1. *Nature*, 382 (1996) 638-642.
29. C.G. Zhang, Z.Q. Jia, B.H. Li, et al., beta-Catenin/TCF/LEF1 can directly regulate phenylephrine-induced cell hypertrophy and Anf transcription in cardiomyocytes. *Biochem Biophys Res Commun*, 390 (2009) 258-62.
30. Z. Ai, A. Fischer, D.C. Spray, et al., Wnt-1 regulation of connexin43 in cardiac myocytes. *J Clin Invest*, 105 (2000) 161-71.

31. T. Toyofuku, Z. Hong, T. Kuzuya, et al., Wnt/frizzled-2 signaling induces aggregation and adhesion among cardiac myocytes by increased cadherin-beta-catenin complex. *J Cell Biol*, 150 (2000) 225-41.
32. C.H. Kim, H. Neiswender, E.J. Baik, et al., Beta-catenin interacts with MyoD and regulates its transcription activity. *Mol Cell Biol*, 28 (2008) 2941-51.
33. M.G. Walker, W. Volkmuth, E. Sprinzak, et al., Prediction of Gene Function by Genome-Scale Expression Analysis: Prostate Cancer-Associated Genes. *Genome Res.*, 9 (1999) 1198-1203.
34. D.J. Pagliarini, S.E. Calvo, B. Chang, et al., A mitochondrial protein compendium elucidates complex I disease biology. *Cell*, 134 (2008) 112-23.
35. J. Lammerding, R.D. Kamm, and R.T. Lee, Mechanotransduction in Cardiac Myocytes. *Ann. N.Y. Acad. Sc.*, 1015 (2004) 53-70.
36. R.S.R.M. Martherus, V.A. Zeijlemaker, and T.A.Y. Ayoubi, Electrical stimulation of primary neonatal rat ventricular cardiomyocytes using pacemakers. *Biotechniques*, 48 (2010) 65-7.
37. R.E. Akins, Jr., D. Rockwood, K.G. Robinson, et al., Three-dimensional culture alters primary cardiac cell phenotype. *Tissue Eng Part A*, 16 (2010) 629-41.
38. B. Ren, F. Robert, J.J. Wyrick, et al., Genome-wide location and function of DNA binding proteins. *Science*, 290 (2000) 2306-9.
39. Mukherjee, S., et al., Rapid analysis of the DNA-binding specificities of transcription factors with DNA microarrays. *Nat Genet*, 2004. 36(12) 1331-9.
40. Jolma, A., et al., Multiplexed massively parallel SELEX for characterization of human transcription factor binding specificities. *Genome Res*, 2010. 20(6) 861-73.
41. N. Aspromonte, D.N. Cruz, R. Valle, et al., Management and monitoring of haemodynamic complications in acute heart failure. *Heart Fail Rev*, (2011).
42. A.M. Cogswell, R.J. Stevens, and D.A. Hood, Properties of skeletal muscle mitochondria isolated from subsarcolemmal and intermyofibrillar regions. *Am J Physiol.*, 264 (1993) C383-389.
43. V. Papalouka, D.A. Arvanitis, E. Vafiadaki, et al., Muscle Lim Protein Interacts with Cofilin 2 and Regulates F-Actin Dynamics in Cardiac and Skeletal Muscle. *Mol Cell Biol.*, 29 (2009) 6046-6058.
44. S.Y. Boateng, S.E. Senyo, L. Qi, et al., Myocyte remodeling in response to hypertrophic stimuli requires nucleocytoplasmic shuttling of muscle LIM protein. *J Mol Cell Cardiol*, 47 (2009) 426-35.

SUMMARY

A key mechanism in determining the spatial and temporal expression of genes is transcription, the process in which RNA intermediates are formed based on their DNA templates. In this dissertation we studied various aspects of transcription in cardiomyocytes, the highly specialized muscle cells that generate the contractile force of the heart. The heart muscle is one of the three types of muscle that are present in the human body. The other two are skeletal and smooth muscle. In **chapter 2** we provided a review of how transcription regulating proteins (transcription factors) contribute to the cardiac-specific expression of genes. In addition we described how these activating and inhibitory factors interact with one another, and eventually form regulatory networks that govern the cardiomyocyte transcriptome. The core network regulating cardiomyocyte transcription seems to be one that is shared with the other muscles. Ubiquitously expressed factors such as SRF connect various elements in all muscle types. MEF2, which acts as a central player in skeletal and smooth muscle, is even so pivotal for transcription of various cardiac-specific genes. Besides the 'shared muscle' factors, however there are some, such as GATA4/6, NKX2-5, HAND2, of which the expression is strongly restricted to the cardiomyocytes. These factors assist in constituting tissue specificity, for example by lifting inhibitory mechanisms which block transcription in other types of cells.

The DNA binding transcription factors target specific sequences. The bulk of these regulatory sequences are located in regions directly 5'-proximal to the genes that they control. For many genes, little to nothing was known about these promoter regions. In **chapter 3** we reported on how we identified promoter regions based on cross-species conservation and occurrence of elements known to be involved in regulating transcription in muscle cells. The genes included in this study were cysteine and glycine-rich protein 3 (*CSRP3*), fatty acid binding protein 3 (*FABP3*), lysyl oxidase-like 1 (*LOXL1*), myosin light chain 3 (*MYL3*), myosin light chain 7 (*MYL7*), heat shock 27 kDa protein family member-7 (*HSPB7*) and titin-capping protein / telethonin (*TCAP*). The selected promoter fragments were cloned in luciferase reporter vectors. The resulting constructs were transfected in cardiomyocytes, skeletal muscle, smooth muscle cells and HELA cells to study their tissue specificity. Many of the selected fragments were shown to display an expression profile similar to that of the endogenous genes. For most of these genes, this study was the first to provide information about their human promoters.

In vivo, the heart is continuously contracting in response to electrical signals. It had been shown that electrical pulsing improve cardiomyocyte function *in vitro*, and that it directly affects the transcription of some heart-specific genes. To enable further studies on the effects of electrical pulses, we developed a culture system in which human pacemakers were employed to continuously provide electrical pulses to the cells in culture. This system is presented in **chapter 4**, together with a refined protocol for culturing neonatal rat ventricular cardiomyocytes (NRVCs).

Using micro-arrays and RT-QPCR we performed a genome wide study on the effect of electric pulsing on the transcriptome of NRVCs in culture (**chapter 5**). By exposing NRVCs to electric pulses in the presence and absence of the contraction inhibiting agent blebbistatin, we ascertained whether genes were responding to the electrical pulses themselves or to the contractions that were triggered by these pulses. Electrical pulsing appeared to activate various processes, including calcium handling, fatty acid metabolism and differentiation. In addition we observed an increased transcription of some genes in the pulsed cultures supplemented with blebbistatin, indicating that stretch negatively affected their expression. In all, we showed that electrical signals play a crucial role in cardiomyocyte development and function.

Genes that are expressed in a similar spatial and temporal manner are prone to harbor a similar set of regulatory sequences in their promoters. In addition, there is often a strong conservation of these regulatory sequences across related species. Based on this, we described in **chapter 6** on a bipartite bioinformatics analysis on the 5' flanking promoter regions of cardiac-specific genes. In the first part, we screened the promoter regions for regulatory motifs and modules involved in obtaining cardiac-specific expression. In the second, we investigated which regulatory elements are involved in the transcriptional response to electric stimuli. Many of the motifs that arose from the search for cardiac-specific elements are shared with the other muscles. Only a very few, like muscle initiator sequence 19, appeared to be more unique to cardiac gene promoters. In the promoters of genes specifically activated by pulsing, we observed an enrichment of binding sites for various transcription factors including E-box 47, GATA, TBX5, MEF2, MYOG and TCF/LEF. By mutating selected promoter elements of MYOG/MyoD, SP1 and LEF1 in a separate luciferase reporter constructs containing the *CSRP3* promoter, we confirmed that all the sites were indeed of crucial importance for its transcriptional activity. In addition, data from our own studies, together with information from literature, led us to believe

that an E-box⁴⁷ element in the MYH6 promoter is involved in the transcriptional response to altered calcium handling, activated by electrical pulsing, while it previously was described mistakenly as being a mediator of the stretch response. Further analysis showed that the LEF1 site in the *CSRP3* promoter is targeted by TCF7L2.

TCF7L2 is known to be acting in concert with β -catenin, a crucial player in Wnt signaling. In our array study we found that some elements of Wnt signaling were activated by pulsing. These two facts combined led us to further investigate the relation between electrical pulsing and Wnt signaling (**chapter 7**). NRVCs were cultured with and without electrical pulsing in different media. The media of separate series were supplemented with either a calcium antagonist or agonist, Wnt3a or Wnt5a conditioned medium, or transiently transfected to obtain over-expression of β -catenin. We found that in the absence of pulsing, the addition of Wnt3a or Wnt5a led to an instantaneous increase in the frequency of contractions. Furthermore, we showed that several cardiac-specific genes respond to electrical pulsing in a similar fashion as in response to an activation of the β -catenin dependent Wnt signaling. Where we showed in chapter 6 that information available on spatial and temporal expression profiles can be used to identify the underlying regulatory motifs and modules, **chapter 8** deals with the question how this information can assist in identifying the functional role of poorly annotated genes. Based on the hypothesis that functionally related genes share a similar expression profile, we set out to identify the functional role of two genes, Coiled-coil-helix-coiled-coil-helix domain containing 10 (*CHCHD10*) and glioblastoma amplified sequence (GBAS). Although both had been identified previously as being involved in the mitochondria, their exact function remained unknown. Based on their expression profiles we predicted that both genes played a role in oxidative phosphorylation. This could be confirmed via experiments in which we silenced the genes. In addition we were able to determine that CHCHD10 plays a role in regulating complex IV activity.

In summary, the major players required for cardiac-specific transcription appear to have been identified, although more study is still required to understand their exact interplay. We found that the cardiac transcriptome is greatly affected by pulsing and identified TCF and E-box 47 as important mediators involved. Furthermore, we showed with the partial elucidation of GBAS and CHCHD10 function, that although many questions still remain in relation to the cardiac transcriptome, answers to other questions can already be found within available transcriptomics data.

SAMENVATTING

Een belangrijk mechanisme bij het bepalen van waar en wanneer genen tot expressie komen is transcriptie, het proces waarbij de RNA tussenproducten worden gevormd op basis van hun DNA-sjablonen. In dit proefschrift hebben we verschillende aspecten van transcriptie onderzocht in cardiomyocyten, de zeer gespecialiseerde spiercellen die de contractiekracht van het hart genereren. De hartspier is een van de drie soorten spieren die aanwezig zijn in het menselijk lichaam. De andere twee zijn de skelet- en gladde spieren. In **hoofdstuk 2** geven we een overzicht van hoe de eiwitten die transcriptie reguleren (transcriptiefactoren), bijdragen aan de hart-specifieke expressie van genen. Daarnaast beschrijven we hoe deze activerende en remmende factoren interageren met elkaar, en uiteindelijk regulerende netwerken vormen die het transcriptoom van de cardiomyocyt controleren. Het kern-netwerk dat cardiomyocyt transcriptie reguleert lijkt te worden gedeeld met de andere spieren. Alomtegenwoordige transcriptie factoren zoals SRF verbinden verschillende elementen met elkaar in alle spiertypes. MEF2, die fungeert als een centrale speler in skelet- en gladde spier, is cruciaal voor de transcriptie van verschillende hart-specifieke genen. Naast de gedeelde spier factoren zijn er enkele, zoals GATA4 / 6, NKX2-5, HAND2, waarvan de expressie sterk beperkt is tot de hartspiercellen. Deze factoren helpen bij het vormen van de weefsel specificiteit, bijvoorbeeld door het opheffen van remmende mechanismen die de transcriptie blokkeren in andere celtypen.

De DNA-bindende transcriptie factoren richten zich op specifieke sequenties. Het grootste deel van deze regulerende sequenties zijn gevestigd in een het deel van het DNA dat direct 5'-proximaal gelegen is van het gen die ze controleren. Voor vele genen is er weinig tot niets bekend over deze promotor regio's. In **hoofdstuk 3** rapporteren we over hoe we promotor regio's geïdentificeerd hebben op basis van de conservering van deze elementen in verschillende diersoorten. De genen die in deze studie werden cysteine en glycine-rijk eiwit 3 (CSR3), vetzuur-bindend eiwit 3 (FABP3), lysyl oxidase-achtige 1 (LOXL1), myosine lichte keten 3 (MYL3), myosine lichte keten 7 (MYL7), *heat shock* 27 kDa eiwit familielid-7 (HSPB7) en titine-kap eiwit / telethonin (TCAP). De geselecteerde promotor fragmenten werden gekloond in luciferase reporter vectoren. De resulterende constructen werden getransfecteerd in cardiomyocyten, de skeletspieren, gladde spiercellen en HELA cellen om hun weefsel specificiteit te analyseren. Veel van de geselecteerde fragmenten vertoonden een expressie profiel vergelijkbaar met die van de endogene genen te geven. Voor de meeste

van deze genen, is deze studie de eerste die informatie verschaft over hun humane promotoren.

In het lichaam trekt het hart voortdurend samen in reactie op elektrische signalen. Het is aangetoond dat elektrische pulsen de functie van cardiomyocyten verbetert *in vitro*, en dat zij rechtstreeks invloed op de transcriptie van een aantal hart-specifieke genen. Om verdere studies naar de effecten van elektrische pulsen, mogelijk te maken, ontwikkelden we een celkweek-systeem waarin pacemakers worden gebruikt om continu elektrische pulsen geven aan de cellen in cultuur. Dit systeem wordt gepresenteerd in **hoofdstuk 4**, samen met een protocol voor het kweken van ventriculaire cardiomyocyten verkregen uit neonatale ratten, (NRVCMs).

Met behulp van micro-arrays en RT-QPCR hebben we een genoom-wijde studie uitgevoerd naar de effecten van elektrische pulsen op het transcriptoom van NRVCMs in cultuur (**hoofdstuk 5**). Door NRVCMs bloot te stellen aan elektrische pulsen in de aanwezigheid en afwezigheid van de contractie-remmende stof blebbistatin, hebben we onderzocht of genen reageren op de elektrische pulsen zelf of aan de contracties die worden geactiveerd door deze pulsen. Elektrische pulsen blijken diverse processen te activeren, waaronder calciumhuishouding, vetzuurmetabolisme en differentiatie. Daarnaast zagen we een toegenomen transcriptie van bepaalde genen in de gepulste culturen waaraan blebbistatin was toegevoegd. Dit geeft aan dat de samentrekking van de cellen hun expressie negatief beïnvloedt. Alles samengenomen tonen we aan dat elektrische signalen een cruciale rol spelen in de ontwikkeling en het functioneren van cardiomyocyten.

Voor genen die op dezelfde tijd en plaats tot expressie komen is het zeer waarschijnlijk dat zijn overeenkomstige regulerende sequenties hebben in hun promotoren. Daarnaast is er vaak een sterke conservering van deze regulerende sequenties in diersoorten met een vergelijkbare biologie. Op basis van hiervan presenteren wij in hoofdstuk 6 een tweeledige bioinformatica analyse van de 5'-flankerende promotor regio van hart-specifieke genen. In het eerste deel hebben we de promotor regio's gescreend voor regulatoire motieven en modules die betrokken zijn bij het tot stand brengen van hart-specifieke expressie. In de tweede hebben we onderzocht welke elementen betrokken zijn bij de transcriptionele respons op elektrische prikkels. Veel van de motieven die naar voren kwamen in de zoektocht naar hart-specifieke elementen worden gedeeld met de andere spieren. Slechts een zeer klein aantal, zoals spier-initiator sequentie 19, lijken meer uniek is voor de promotoren van de hart-specifieke genen. In de

promotoren van genen die specifiek geactiveerd worden door pulseren, zagen we een verrijking van bindingsplaatsen voor diverse transcriptie factoren, waaronder de E-box 47, GATA, TBX5, MEF2, MYOG en TCF/LEF. Door het muteren van geselecteerde promotor elementen van MYOG / MyoD, SP1 en LEF1 in aparte luciferase reporter constructen met de CSRP3 promotor, hebben we bevestigd dat deze elementen inderdaad van cruciaal belang zijn voor transcriptie. De resultaten van onze eigen studies, samen met de informatie uit de literatuur, impliceren dat een E-box47 element in de MYH6 promotor inderdaad betrokken is bij de transcriptionele respons op veranderde calciumhuishouding, geactiveerd door de elektrische pulsen, terwijl deze eerder ten onrechte werd omschreven als effector van de reactie op contractie. Verdere analyse toonde aan dat de LEF1 site in de CSRP3 promotor een bindingsplaats is voor TCF7L2.

Van TCF7L2 is bekend dat het, in samenwerking met beta-catenine, een cruciale speler is in Wnt-signalering. In onze array-studie vonden we dat sommige onderdelen van Wnt-signalering worden geactiveerd door pulseren. De combinatie van deze twee feiten leidde ons ertoe om verder onderzoek te doen naar de relatie tussen elektrische pulsen en Wnt-signalering (**hoofdstuk 7**). NRVCs werden gekweekt, met en zonder elektrische pulsen in verschillende kweekmedia. De media van aparte serie werden aangevuld met ofwel een calciumantagonist of agonist, Wnt3a of Wnt5a geconditioneerd medium, of transient getransfecteerd om over-expressie van beta-catenine te verkrijgen. We vonden dat in de afwezigheid van pulserende, de toevoeging van Wnt3a of Wnt5a leiden tot een onmiddellijke stijging van de frequentie van de contracties. Verder toonden we aan dat de reactie van verschillende hart-specifieke genen op elektrische pulsen sterke gelijkenis vertoont met de reactie die optreedt wanneer de beta-catenine afhankelijke Wnt-signalering wordt geactiveerd.

In hoofdstuk 6 hebben we laten zien dat de beschikbare informatie over ruimtelijke en temporele expressie profielen gebruikt kunnen worden voor de identificatie van de onderliggende regulatoire motieven en modules. In **hoofdstuk 8** demonstreren we hoe deze informatie gebruikt kan worden bij het identificeren van de functionele rol van slecht geannoteerde genen.

Gebaseerd op de hypothese dat functioneel gerelateerde genen een soortgelijke expressie profiel te delen, hebben wij een studie ondernomen naar de functionele rol te identificeren van twee genen:

coiled-coil-helix-coiled-coil-helix domain containing 10 (CHCHD10) en glioblastoma geamplificeerde sequentie (GBAS). Hoewel beide al eerder waren geïdentificeerd als zijnde betrokken bij de mitochondriën, hun precieze functie is onbekend gebleven. Op basis van de expressie profielen voorspelden we dat beide genen een rol speelden in de oxidatieve fosforylering. We hebben dit kunnen bevestigen via experimenten waarin we de producten van genen, het mRNA, inactiverden. Bovendien waren we in staat om dat te bepalen CHCHD10 een rol speelt in het reguleren van complex IV activiteit.

Samengevat, de belangrijkste transcriptie factoren die nodig zijn voor hartspecifieke transcriptie lijken te zijn geïdentificeerd. Echter, meer studies zijn nog nodig om te begrijpen hoe het samenspel tussen deze factoren exact verloopt. We vonden dat het transcriptoom van het hart sterk wordt beïnvloed door elektrisch pulseren en we identificeerden TCF en E-box 47 als belangrijke mediators voor deze respons. Verder demonstreerden in onze studie naar de functie van GBAS en CHCHD10, dat bestaande publiek toegankelijke transcriptomics data gebruikt kan worden om nieuwe onderzoeksvragen te beantwoorden.

DANKWOORD

Na al die pipeteer-sessies, de vele agarose-gelletjes, talrijke contracties van mijn eigen cardiomyocyten *in vivo* en die van ratten *in vitro*, en een afschrikwekkende hoeveelheid van typewerk en daaruit voortvloeiende computer-gerelateerde frustraties en 3 laptops later, is het dan eindelijk zover: het proefschrift is daar! Ja, niet mijn proefschrift maar hét proefschrift, immers dit epistel zou er niet zijn gekomen zonder enige vorm van medewerking van een aantal mensen. Deze wil ik hier nog 'even' mijn speciale dank toekennen.

Als eerste Torik. Dit proefschrift is voortgekomen uit jouw geestekind: een onderzoeksproject voor het bestuderen van cardiovasculaire promotoren. Uiteindelijk is een en ander een beetje anders uitgevallen dan we bij de start hadden kunnen vermoeden, maar dat getuigd ook van de rijkdom aan ideeën die jij hebt. De verschillende brainstormingen leverden genoeg inspiratie op om desnoods nog een flink aantal proefschriften te vullen. Nu het boekje er ligt ervaar ik nog steeds een nasmaak van transcriptie, eentje die smaakt naar meer! Bedankt voor de fijne samenwerking en het delen van je kennis van de wetenschap, pepers (geen typefout!) en cocktails.

Bert, een groot deel van mijn werkzaamheden had betrekking op andere zaken dan mito's. Desondanks werd mijn verblijf in Popgen/Clinical Genomics toch gedoogd, en ben je ook nog eens promotor. Mede dankzij jouw commentaren op artikelen en aanvullingen op het onderzoeksbudget heb ik het een en ander tot een goed einde kunnen brengen. Bedankt voor dit alles en zeker ook voor de hulp bij de laatste loodjes van het gehele promotie-traject.

Joep, ik heb veel geleerd van jouw commentaren in de eerste fasen van mijn schrijfwerk van artikelen. En, ook al leverden de suggesties voor de verbeteringen van dit boekje flink wat extra type-werk op, ik heb deze toch zeker kunnen waarderen! Ervaring bleek maar weer eens van ongekende waarde te zijn.

Erika en Bieke, mijn trouwe lab-maatjes. Op het lab hebben jullie het flink te verduren gehad. Sorry als ik jullie levensduur verkort blijkt te hebben als ik weer eens onaangekondigd en (voor jullie) plotseling mijn entree maakte. Bedankt allebei voor jullie hulp bij het kweken van de vele celletjes, het slopen van ze, de arrays, het pipeteren van die afgrijselijke 384 well platen en al die andere zaken. En, Erika, oude pipeteer-beer, dankjewel dat je mij zo goed bekend hebt gemaakt met

het vocabulair dat regulier gebezigd wordt tussen de gebruikers van het lab. Er rust jouw denk ik nog een schone taak om de rest van de Universiteit Maastricht hierin te onderwijzen. Dit gezien de optrekkende wenkbrauwen van menig bezoeker. Ik ben blij dat ik het geheel kan afsluiten met jullie als paranimfen aan mijn zijde.

Veel dank aan de mensen van Leuven. Wouter van Delm, Yves Moreau en Peter van Loo. Jullie hebben een groot deel van dit werk mede mogelijk gemaakt, super veel dank hiervoor. En ook voor jullie geduld en kunde om de geavanceerde bioinformatica te vertalen in een vorm die mensen zoals ik kunnen begrijpen.

A special thanks to Gordana Vunjak Novakovic and Nina Tandon at Columbia University, New York City. Thank you for allowing me to spend some time in your lab and learn from your expertise on culturing cardiomyocytes. The experiences I was able to pick up in NYC were invaluable to completing the research presented here.

Volkert Zeilemaker, ik zou niet weten hoe dit boekje er had uitgezien zonder de pacemakers en niet te vergeten ook jouw interesse voor ons onderzoek. Ik ben blij dat wij hebben kunnen delen in jouw rijkdom aan ervaring met deze schokkende apparaten.

Wim Sluiter, jij hebt maar weer eens bewezen dat kant en klaar kitjes geen vervanging zijn voor echte expertise. Dankzij jouw analyses en onze gedachtenwisselingen via e-mail en 'snail-mail' is het mito verhaal toch wat completer geworden.

Hoewel plaatjes schieten een van mijn hobbies is... zodra er nog een confocaal tussen zit wordt het toch wel andere koek. En filmpjes maken van die hyperactieve cellen is ook een kunst. Veel dank aan Jos Broers voor de hulp bij dit alles.

Hoewel later nog wel een keer aan de beurt, hier alvast mijn wetenschappelijke dank aan Hilde. Dank je voor de fijne samenwerking. Het Wnt-verhaal wat toch een van de sluitstukken is geworden, was er zonder jouw bijdrage niet geweest.

Dan..., mijn 'lab-slaven': Tom, Ronald, Cris en Lars. Slaaf in deze context staat (nog) niet in de Van Dale. Maar ja, een beetje V.O.C. mentaliteit op het lab ook niet verkeerd. Beste 'lab-slaven'/gewaardeerde

medewerkers, jullie hebben het zwaar te verduren gehad, maar weet dat jullie harde werk echt gewaardeerd is en zeker ook heeft bijgedragen aan dit proefschrift. Ook jullie niet-wetenschappelijke bijdrages heb ik zeker kunnen waarderen. Iedere keer als ik een Kinder-surprise ei in handen krijg zal ik weer met weemoed aan jullie terugdenken.

Tijd voor mijn kamer genoten. Eerst, de 'oude': Bianca (vrouwtje mito), Dr. Dr. Lars, Rudy en Nicole. Bedankt voor jullie gezelschap op de kamer. Jullie al dan wel of niet wijze AIO-levenslessen worden nog immer doorgegeven aan de nieuwe generaties van wetenschappers.

Merci aan de dames die menige krachtterm en ook de vele mislukte lanceerpogingen van mijn laptop hebben moeten aanhoren. Florence, tja toch eerder!, An, toonbeeld van integratie in Holland! Rita, thanks for presenting me with a target for aiming my shuttle at, and introducing me to the tasty side of Portugal. I already started to miss it, especially the flamed chorizo and liquor shots in chocolate (served separately of course!).

Mike, een gelijkgestemde in de buurt was soms wel eens fijn, voor ons dan. Zeker wanneer je onder hypoxische en oestrogeen-geïnfesteeerde condities moet werken. Ik vertrouw erop dat je mijn gedachtegoed in leven houdt. En als je nog eens wilt oefenen in Nederlands spreken, je weet me te vinden!

Mijn overige (voormalige) popgen/clinical genomics collega's: Fons en Miroslav (daanke voor de hulp met jullie java tooltjes). Patrick, Jo, Ton bedankt voor de af en toe broodnodige IT support. Wanwisa, (B)Iris, Frank, Ellen, Marion, Sabine, Jos, Rosy, Roselie, Marij, Caroline. Merci voor alle bijdrages in o.a. de gezelligheid in het lab en er buiten.

Ook al was het soms niet te merken, gelukkig bleek er nog een leven te zijn buiten het lab. Daarvoor wil ik ook hier nog even een paar mensen bedanken.

Hilde, Bea en Erik, Ik heb veel genoten van de samenkomsten van ons kwartet. Biertjes, wijntjes en diners op het Vrijthof, bij iemand thuis tot zelfs in New York City, Daytona Beach en in het Mount Everest Basecamp in Tibet. Samen delen in de ervaringen in de wetenschap, zonsopgang bij de Atlantische oceaan, (smokey) whisky's proeven bij Sue en de spaceshuttle Atlantis omhoog zien gaan. Allemaal speciale

momenten. Ik ben blij dat ik die met even zo speciale mensen heb mogen delen. En Erik, bedankt voor het badminton gezelschap, ik denk met veel plezier terug aan de vele duels die wij tegen en ook met elkaar hebben uitgevochten op het slagveld. Ook als we allemaal Dr. zijn moeten we met zijn vieren periodiek maar blijven evalueren of dat toevluchtsoord in de Stille Zuidzee er toch moet komen.

Nu we toch in de feesten en partijen zijn beland. Natuurlijk is er nog 'de club' Marion, Nuria, Esther en Mirjam, merci voor de gezelligheid, de geestelijke bijstand, de hulp bij het ledigen van mijn cocktail collectie in Meendaal 124B, de sportieve en zeker ook gezellige momenten in de sporthal, de films in Lumiere, de drankjes op het Vrijthof, de wandelingen door de bossen, de chocolademelk/koffie sessies in de koffie corner, de bezoeken aan de mergelgroeven, vuursteenmijn, etc etc...

Hoewel Maastricht een leuke stad is en veel te bieden heeft, blijft een andere veilige thuishaven altijd van belang, zeker voor een 'Hollander' die niet veel van carnaval moet hebben. Bedankt dus aan mijn ouders bij wie ik altijd kan aankloppen op welk moment dan ook. Ook al zijn mijn activiteiten vaak voor jullie letterlijk een "ver het bed show". Zonder jullie was ik nooit zover gekomen dus de eer van het bereiken van deze mijlpaal valt ook zeker jullie ten deel. Bedankt voor alles uit het verleden en alvast ook voor alles in de toekomst.

Tevens een woord van dank naar de rest van het thuisfront zusjes en (schoon)-broer, bedankt voor jullie support van de afgelopen jaren.

Klein van stuk maar van groots betekenis, zo ook mijn woorden:
Mercies voor alles.

CURRICULUM VITAE

Sarkies Ruben Mattheus Martherus werd geboren in Eemnes op 30 oktober 1977. Na het behalen van het diploma van het Hoger Algemeen Voortgezet Onderwijs in 1996 aan de Scholengemeenschap Werenfridus te Hoorn doorliep hij het Hogere Laboratorium Onderwijs aan de Hogeschool van Amsterdam, en ontving daar in 2000 zijn diploma van de opleiding biotechnologie. Vervolgens volgde hij aan de Universiteit van Wageningen het internationale Master of Science programma wederom met de specialisatie biotechnologie. Direct na het behalen van zijn Masters diploma in 2002 vertrok hij voor 15 maanden naar Wuhan in de Volksrepubliek China alwaar hij als onderzoeker in dienst was van het Laboratorim voor Moleculaire Virologie van de Chinese Academie van Wetenschappen. Na een jaar terug in Nederland waarin hij als uitzendkracht werkte begon hij in januari 2005 met zijn onderzoek aan de Universiteit van Maastricht. Verbonden aan de vakgroep Clinical Genomics verrichte hij daar onder leiding van Dr. Torik Ayoubi onderzoek uit naar verschillende aspecten van transcriptomics in relatie tot het hart. Dit onderzoek wordt gepresenteerd in dit proefschrift en delen hiervan zijn gepubliceerd in internationale wetenschappelijke literatuur.

Sinds maart 2010 is de auteur werkzaam als Post-Doc onder leiding van Prof. Dr. Jeffrey Towbin bij het Heart Institute van Cincinnati Children's Hospital in Cincinnati, Ohio in de Verenigde Staten van Amerika.

LIST OF PUBLICATIONS

Martherus RS, Sluiter W, Timmer ED, VanHerle SJ, Smeets HJ, Ayoubi TA.

Functional annotation of heart enriched mitochondrial genes GBAS and CHCHD10 through guilt by association.

Biochem Biophys Res Commun. 2010 Nov 12;402(2):203-8.

Martherus RS, Vanherle SJ, Timmer ED, Zeijlemaker VA, Broers JL, Smeets HJ, Geraedts JP, Ayoubi TA.

Electrical signals affect the cardiomyocyte transcriptome independently of contraction.

Physiol Genomics. 2010 Nov 29;42A(4):283-9.

Martherus RS, Zeijlemaker VA, Ayoubi TA.

Electrical stimulation of primary neonatal rat ventricular cardiomyocytes using pacemakers.

Biotechniques. 2010 Jan;48(1):65-7.